

Grammatical Relations Typology

Balthasar Bickel
University of Leipzig

1. Grammatical relations past and present

Traditionally, the term ‘grammatical relation’ (GR) refers to the morphosyntactic properties that relate an argument to a clause, as, for example, its subject or its object. Alternative terms are ‘syntactic function’ or ‘syntactic role’, and they highlight the fact that GRs are defined by the way in which arguments are integrated syntactically into a clause, i.e. by functioning as subject, object etc. Whatever terminology one prefers, what is crucial about the traditional notion of GRs is (a) that they are identified by syntactic properties, and (b) that they relate an argument to the clause.¹ This differentiates GRs from semantic roles (SRs), also known as thematic roles (θ -roles): SRs are semantic, not syntactic relations, and they hold between arguments and predicates (typically verbs), rather than between arguments and clauses. The difference between GRs and SRs is best visible in such contrasts as *Sue has killed the shark* vs. *Sue was killed by the shark*: in both cases, the NP *Sue* is the subject of the clause. But in the active version, the referent of *Sue* is the agent of ‘kill’, while in the passive version, *Sue* is the patient of ‘kill.’

The syntactic properties that have traditionally been considered the key identifiers of GRs are the property of triggering verb agreement and the property of being assigned a specific case. In our example, *Sue* triggers third person singular agreement in the verb and this identifies the NP as the subject of the clause. In some languages, e.g. Russian or Turkish, the subject would furthermore be identified by nominative case assignment.

Research over the past three decades has greatly expanded the range of syntactic properties that identify GRs in particular languages, and one of the most important results of this research is that properties often do not converge on a single set of GRs in a language. Consider the following examples from Nepali:²

(1) Nepali (Indo-European; Himalayas)

- a. *ma* *ga-ẽ*.
1sNOM go-1sPST
‘I went.’

¹ Technically, adjuncts also bear grammatical relations in this sense. In this chapter I only discuss argument relations; for some suggestion on how adjuncts can align with arguments in grammatical relations, see Bickel and Nichols in press.

² Glossing follows the Leipzig Glossing Rules (<http://www.eva.mpg.de/lingua/files/morpheme.html>), with the following additions: ACT ‘active’, AUG ‘augmented (number)’, c ‘complementary number (Kiowa)’, CNT ‘continuative’, CONTR ‘contrastive’, DIR ‘direct’, DISCNT ‘discontinuative’, DO ‘direct object’, DS ‘different subject’, INTS ‘intensifier’, INV ‘inverse’, LNK ‘linker’, MONO ‘monofocal = 1st person or singular non-first person’, NML ‘nominal (marking free stems)’, OBV ‘obviative’, PAT ‘patientive’, PROX ‘proximative’, PUNCT ‘punctual’, REM ‘remote’, SAP ‘speech-act participant’, SEQ ‘sequential’, SIM ‘simultaneous’, SS ‘same subject’, SUB ‘subordinator’, V ‘verb derivative’.

b. *mai-le timro ghar dekh-ẽ.*
 1s-ERG your house.NOM see-1sPST

‘I saw your house.’

In both examples, the expression for ‘I’ is identified as the subject of the clause by the fact that it triggers verb agreement (–ẽ ‘first person singular past’). But with regard to case, we are dealing with two different relations: the one identified by nominative case (*ma*) in (1a), and the one identified by ergative case (*maile*) in (1b). Examples like these multiply in many ways when we expand our dataset of languages across the world, and even more when we look, as we will do in this chapter, at the syntactic properties of arguments beyond agreement and case, e.g. at the behavior in relative clauses or raising constructions. This finding has become known in the literature as the construction-specific nature of grammatical relations.

The construction-specific nature of GRs poses important problems for the traditional view. As noted above, an argument is traditionally said to bear a GR to a clause, and properties like case and agreement serve as ‘diagnostics’ or ‘tests’ for identifying the GRs of the clause. A first problem with this view is that, as we just saw, these tests often do not converge on the same GR. This makes it unclear which GR is borne by an expression like ‘I’ in (1): is it, or is it not the subject of the clause? If we say ‘yes’ on the account that ‘I’ triggers agreement in both examples, what do we gain beyond replacing the term ‘agreement-trigger’ by the term ‘subject’, and how do we explain the fact that the same GR gets different case marking in the two examples? Moreover, given that *ma* ‘I’ in (1a) bears the same case as *timro ghar* ‘your house’ in (1b), shouldn’t we rather say that this argument, the one in the nominative, is the subject? But then, why go with the evidence from case marking rather than from agreement? A second problem is that the traditional view treats properties like agreement as if they were test tools for the linguist rather than grammatical devices in their own right but this deflects from the crucial question why some devices seem to define GRs in one way while others define them in other ways. For example, there are many languages like Nepali where agreement treats ‘I’ the same way in the two examples, while case differs — but only few languages that do it the other way around (cf. Siewierska 2004:53f): where agreement differs in the two examples, while case is the same. Why? As long as case and agreement are seen merely as diagnostic identifiers, the question is even difficult to ask.

The properties that define GRs receive their deserved center-stage status as soon as we reconceptualize the notion of GR as the syntactic relation that an argument bears TO A SPECIFIC CONSTRUCTION OR RULE³ rather than to the clause in which the argument is realized. Thus, in (1), ‘I’ bears one GR to the agreement construction (the same in a and b), and one GR to the case construction used (not the same in a and b). In general, then, a GR is defined as the set of arguments that is selected by a construction for a particular syntactic purpose, such as, for example, for agreement rules or case government. This means that an argument can bear as many GRs as it enters

³ While the difference between rules, constructions and ordered constraint sets is of critical importance for the architecture of formal grammar models, it is irrelevant for defining typological variables. Also, it is irrelevant for typology whether GRs are mathematically modeled as equation matrices (like, e.g., in LFG or Construction Grammar) or as graph-theoretical nodes (like, e.g., in Minimalism). All that matters for qualitative typology is that phenomena like case marking or agreement can be precisely identified across languages, that they can be coded as for how they involve GRs, what arguments these GRs include, and what other grammatical properties the phenomena have.

constructions in a given syntactic context, and these GRs need not be the same across constructions. How GRs are selected, how they are defined for each construction, to what degree their distribution overlaps across constructions, how types of GRs correlate with each other and how they distribute in the languages of the world — these are the core issues that define research in GR typology.

In the following, I first review the typological variables that define or condition specific GRs (Sections 2 and 3). In Section 4 I survey the kinds of constructions that have GRs and in Section 5 I look into interactions between GR definitions in different constructions. Section 6 briefly addresses issues of worldwide distributions and Section 7 concludes with suggestions for future research.

2. Defining grammatical relations

GRs are equivalence sets of arguments, treated the same way by some construction in a language, e.g. being assigned the same case in a language, or triggering the same kind of agreement. Arguments in turn are defined — to take up Evans's (1997) apt simile — by cast and role: each predicate takes a cast of characters, and each member of the cast plays a distinct role. In more technical terms, arguments are defined by both their relation to the predicate (as agent, theme, etc.) and their referential type (as animate, speaker, topic etc.). Languages vary as to whether their GRs select arguments on the basis of role or reference properties or by combining these two kinds of properties. In the following, I first discuss relational roles, then referential properties of arguments (to which I devote more space since these are less well-known).

2.1 Roles

Arguments bear specific semantic relations to the predicate, e.g. 'the one who sees', 'the one who sleeps', 'the one who is given sth.', etc. A very successful theoretical proposal, which I will follow here, is that for the purpose of GR specifications, such individual, predicate-specific roles merge systematically into generalized roles, sometimes called macroroles (Foley and Van Valin 1984; Van Valin and LaPolla 1997; Van Valin 2005) or protoroles (Dowty 1991; Primus 1999, this volume). This reduces the range of predicate-specific roles into a small set of generalized argument roles that are referenced by specific constructions.

There are various ways to define this set, but the theory that has proven to best capture typological variance is one that defines the set as the minimal set distinguished by numerical valence, i.e. by the distinction between intransitive (one-place), transitive (two-place) and ditransitive (three-place) verbs.⁴ This leads to the by-now classic schema of labels introduced by Comrie (1978) and Dixon (1979a): S 'sole argument of an intransitive verb', A 'most actor-like argument in a transitive verb', and O 'not most actor-like argument in a transitive verb' (Comrie actually uses P here). In order to further distinguish among the two non-actor-like arguments of ditransitives, I will use G for the most goal-like or ground-like (e.g. the one who is given something, or the one to which something is applied) and T for the other (most patient-like) argument (e.g. that which is given or that which is applied to something); cf. Figure 1.

⁴ For the sake of space, I disregard four-place predicates like causatives or benefactives of ditransitives, although in some languages they are an important class that deserves more attention than the issue has traditionally received.

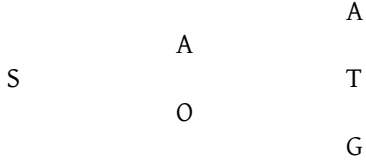


Figure 1: Numerical valence and generalized argument roles

Note that if the set is defined, as it is here, by the minimal distinctions required by the three basic numerical valences, we also expect languages to distinguish between the A of transitives ('A1') and the A of ditransitives ('A2'). This expectation is met by Gyarong, where case marking is sensitive to the distinction between A1 and A2:

(2) Gyarong (lCog-rtse rGyal-roñ) (Sino-Tibetan; Himalayas; Nagano 1984)

- a. *nəyo-ki chigyo kəw-nasño-ch ko.*
 2s-ERG (A1) 1d.NOM (O) 2>1-scold-1d AUX

'You (s) scold us (d).'

- b. *nəyo chigyo kəw-wu-ch ko.*
 2sNOM (A2) 1dNOM (G) 2>1-give-1d AUX

'You (s) give (it to) us (d).'

The sentence in (2a) is monotransitive and its A argument is obligatorily marked by the ergative in *-ki*. With ditransitive verbs like 'give', by contrast, no such marking occurs on the A argument.⁵ Such distinct treatment of A1 and A2 is rare, presumably a result of the great overlap between A1 and A2 in semantics. In the following, A1 and A2 will be subsumed under the cover-term 'A'.

In the simplest case, the generalized argument roles defined in Figure 1 suffice to define the GRs in a given language as specific subsets, in the limiting case as a subset with one member, e.g. a GR allowing only S arguments. The most frequent GRs so defined are given in Table 1. They are also (beginning with Plank 1979) called 'alignment types', a term that highlights the fact that by subsetting arguments, they are being aligned which each other so that they can receive the same treatment by a specific construction, e.g. so that they can all trigger the same agreement paradigm on the verb, or so that they can all be assigned the same case marking.

<i>Grammatical relation</i>	<i>Commonly used names</i>
{S}	intransitive subject, nominative
{S,A}	subject, nominative; accusative alignment
{A}	transitive subject, ergative
{O,T}	direct object, accusative; indirective alignment
{O,G}	primary object, dative; secundative alignment
{T}	secondary object
{G}	indirect object, dative
{S,O,T}	absolutive; nominative; ergative alignment
{S,O,G}	absolutive; nominative; ergative alignment

Table 1: Some common GRs defined as subsets of generalized argument roles

⁵ As important as this observation is for appreciating the true range of typological variation, the distinction between A1 and A2 is currently becoming lost among younger speakers of Gyarong, probably because of increased exposure to Chinese (Nagano, personal communication, October 2003).

The terminology for the GRs in Table 1 is heterogeneous, and when comparing different languages, it sometimes helps to avoid ambiguous terms like ‘subject’ or ‘object’ and use instead names that directly refer to the defining properties of the GR, e.g. ‘the {S,A}-relation’, or ‘the {O,T}-relation’. But occasionally the traditional terms are also useful, and I sometimes use ‘subject’ for {S,A} relations (following Dixon 1994) and ‘object’ for any relation that contains at least O (and perhaps also S or T or G). An additional term that is frequently used for some GRs is the term ‘pivot’, popularized by Dixon (1979a) and Foley & Van Valin (1984). This term refers to any of the subsets in Table 1 but is limited to the special case of a GR in a biclausal construction, e.g. the GR that is referenced in some languages by switch-reference constructions (cf. Section 4.6). For some other biclausal constructions, e.g. control and raising constructions (cf. Section 4.5), the terms ‘controller’ and ‘controllee’ are useful.

2.2 Reference

Of the GRs listed in Table 1 we already encountered the {S,A}, the {A}, and the {S,O,T} relation in the Nepali example (1). The {S,A} relation is instantiated by the agreement construction: only S and A arguments trigger agreement. The {S,O,T} relation is referenced by nominative case, which is in opposition to the {A}-marking ergative. However, this alignment only holds as long as the O argument is inanimate, as it did in (1b). If it is animate or otherwise socially important (‘O-high’), it receives the same dative marking that is also generally used in Nepali for the G argument of ditransitives (Pokharel 2054):

(3) Nepali

- a. *mai-le Prembahādur-lāi dekh-ē.*
 1s-ERG P.-DAT (O) see-1sPST
 ‘I saw Prem Bahadur.’
- b. *mai-le celi vahā-haru-lāi di-ē.*
 1s-ERG (A) marriageable.female.clan.relative.NOM (T) 3h-p-DAT (G) give-1sPST
 ‘I gave them a *celi* (in marriage).’

Because the O argument in the monotransitive sentence in (3a) is animate it is aligned with the G argument of ditransitives, as illustrated by (3b), and receives the same dative case marker. This yields a dative-bearing {O-high, G}-relation. This is then in double opposition to both the ergative {A}-relation and the nominative {S, O-low, T}-relation, resulting in what is called tripartite alignment. (Note that the T argument always remains in the nominative, even when it is human, as in (3b)).

The phenomenon we have just looked at in Nepali is also known as differential object marking (beginning with Bosson 1985): O arguments are mapped into different GRs (notably, direct vs. primary object) for some construction, depending, mostly in a probabilistic rather than categorical way, on such referential properties as animacy, humanness, definiteness, specificity or more general notions of saliency.⁶

⁶ Therefore, an adequate understanding of the way referential properties affect GR choice requires statistical analysis, e.g. multiple logistic regression as proposed by Bresnan et al. 2004. Unfortunately, for most languages we lack corpora of adequate size so that many statements must remain impressionistic

The constructions mostly affected by such referential conditions are case and agreement construction. Nepali was an example with case. An example with agreement comes from Swahili:

(4) Swahili (Benue-Congo; East Africa; Seidl and Dimitriadis 1997)

mbwa a-li-ona mbuzi. a-li-kata kamba na ku-kimbia. a-li-m-rarua
 dog 3s-PST-see goat 3s-PST-cut rope and INF-run 3s-PST-3sO-tear.apart
mbuzi vipande vipande.
 goat part part

‘The dog saw a goat. It cut the rope and ran free. It tore the goat to pieces.’

Mapping O arguments into the agreement-triggering object relation is more likely if the referent is animate (especially human) and/or known to the hearer. Thus in (4), the patient *mbuzi* ‘goat’ only achieves full objecthood in the last clause where the referent is already known and established. In the initial clause, *mbuzi* is new and therefore projected into a different kind of object, one that does not trigger agreement but is in all other respects the same as the general object relation.

Two related responses to O arguments that do not make it into the regular object relation because of their ‘low’ referential status are antipassivization and incorporation. The difference between these and agreement-dropping as in Swahili is a matter of degree, and depends on the number of constructions in which the O argument still behaves as a direct or primary object: in canonical agreement-dropping, the O argument behaves like an object in all constructions except the agreement construction itself; in canonical antipassives and under incorporation, the O argument does not behave like an object in any construction — but there are many cases in between, cf. e.g. Bickel et al. (2007) for a recent case study.

Yup’ik Eskimo illustrates a canonical antipassive: indefinite, non-specific or mass-noun Os are usually not treated as objects for the purpose of agreement (like in Swahili) and also for the purposes of case-marking (where they receive ablative instead of absolutive case) (cf. Mithun 1999: 234f, 408). In the following data from Central Yup’ik Eskimo, (5a) is an active sentence, where the O argument (*nutek* ‘gun’) is definite and therefore functions as an object. As such, it is marked by the absolutive case and triggers agreement. In (5b), by contrast, the O argument is indefinite and can therefore not be an object. As a result it appears as an oblique NP in the ablative, and the verb is marked as antipassive:

(5) Central Yup’ik Eskimo (Eskimo-Aleutan; Alaska; Reed 1977)

- a. *angute-m tamar-a-a nutek.*
 man-sERG loose-TR-3s>3s gun.sABS
 ‘The man loses the gun.’
- b. *angun tamar-i-u-q nuteg-mek.*
 man.sABS loose-ANTIP-INTR-3sS gun-sABL
 ‘The man loses a gun.’

Chukchi also has antipassives, but they are rarely used for regular main clause purposes. The most prominent response to non-salient O arguments in this language is incorporation:

(6) Chukchi (Chukchi-Kamchatkan; Siberia; Dunn 1999)

<i>taŋ-amənan</i>	<i>Cəkwaŋaqaj</i>	<i>ya-qora-nm-at-len.</i>	<i>qora-ŋə</i>	<i>təm-nen</i>
INTS-alone	C.3sABS	PRF-reindeer-kill-V-3s	reindeer-3sABS	kill-3s>3s
<i>ŋelyə-n</i>	<i>jən-nen</i>			
hide-3sABS	take.off-3s>3s			

‘Cakwangaqaj all by himself slaughtered reindeer. He killed a deer [and] took off its hide.’

The first clause describes the activity of reindeer-slaughtering, with no reference to any specific O referent. In the following clause, by contrast, the speaker refers to a specific reindeer (cf. the NP *qoraŋə*), which is also the topic of subsequent clauses. With this referential status, the O argument is now treated as a full-fledged object, and appears as an independent NP, bearing absolutive case and triggering object agreement.

Referential properties are also important for the mapping of A arguments. A fair number of languages allows arguments in the {S,A} relation only if they are animate and/or topical — a pattern that is sometimes called ‘differential subject marking’. For case constructions, this can again be illustrated by Nepali. While in past tense contexts (see examples (1) and (3) above and the discussion in Section 3.3 below), A arguments are always in the ergative-marked {A}-relation, in nonpast contexts A arguments are regularly included in the nominative {S,A}-relation. However, the odds for this {S,A}-status decrease if the A is abstract or inanimate or non-topical; in all these cases, A arguments are more likely to be projected into an ergative {A}-relation even in the nonpast (Clark 1989; Pokharel 2054):

(7) Nepali

- a. *mero sāthi momo khāi-rahe-cha.*
 my friend.NOM Tibetan.dumplings.NOM eat-IPFV-3sNPST
 ‘My friend is eating momos.’
- b. *dhumrapān-le aru-lāi kharab gar-cha.*
 smoking-ERG other-DAT harm do-3sNPST
 ‘Smoking harms others.’
- c. *bāhira ke-ko khalbal? — karmi-haru-le chānā hāli-rahe-cha.*
 outside what-GEN noise worker-p-ERG roof.NOM lay-IPFV-3NPST
 ‘What’s the noise about outside? — It’s the workmen laying the roof.’

The A argument in (7a) is animate and topical, and it is therefore mapped into a nominative-marked {S, A-high} relation. But in (7b) and (7c), the A arguments are assigned the ergative-bearing {A}-relation: in (7b) because the A is inanimate and in (7c) because it is focal.

In some languages, the odds for inanimate A arguments to function as {S,A} are virtually zero. If there is no competing {A} relation available (as there is in Nepali), the response to this constraint is parallel to the treatment of indefinite or inanimate

Os in such languages as Eskimo and Chuckchi: the inanimate A is demoted by diathesis, or it is incorporated. Kiowa has both options:

(8) Kiowa (Kiowa-Tanoan; Eastern North America; Watkins and McKenzie 1984)

a. *tê:-gyà é-thêm.

ice-NML 3sA>3cP-break.PFV

Intended: 'The ice broke it.'

a'. têt:-gyà phí: nò ýhà-dè è-thém-gyá.

ice-NML [3sS]-heavy and.DS there-DIR 3cS-break-DETRANSITIVE.PFV

'The ice_i is heavy, and therefore it_{ij} got broken.'

b. *góm-gyà é-thêm.

wind-NML 3sA>3cP-break.PFV

Intended: 'The wind broke it.'

b'. è-góm-thém-gyá.

3cS-wind-break-DETRANSITIVE.PFV

'The wind broke it.' (Literally, 'it got wind-broken.')

Using A arguments in the agreement-triggering subject function (8a,b) is ungrammatical. The grammatical version in (8a') deletes the A argument of 'break', and the verb appears in a detransitized form functioning as a passive. In (8b'), the A argument is incorporated into the verb. Both options effectively block the A argument from bearing any GR for any construction in the language.

Referential properties are most central to what are commonly called hierarchical systems. In all examples discussed so far, GRs are defined by pre-selecting one of the transitive arguments (A or O) to combine with the S argument or by preselecting one of the ditransitive arguments (T or G) to combine with the O argument. In hierarchically-defined GRs, by contrast, all, or nearly all arguments compete for the same GR, and the choice among arguments rests on referential properties alone.

An often-discussed example of this is what one finds in some Austronesian languages (especially those in the Philippines and those in Taiwan). In each clause, one NP is selected as the principal GR, variously identified in the literature as 'topic', 'focus', 'pivot', 'nominative' or 'subject'. I will use the term 'proximative' because all other terms have well-established uses at odds with the nature of the principal GR under hierarchical alignment. The proximative GR is marked by *ang=* in Tagalog and is referenced by a number of constructions, e.g. by conjunction reduction, relative constructions or floated quantifiers (see Sections 4.3, 4.7 and 4.8, respectively). The choice of which NP bears the proximative GR depends exclusively on referential properties and can fall on any argument (S, A, O, T, or G) or adjunct: all that matters is that the NP has specific reference and that it is the most topical element in discourse (indicated here by italics in the translation):

(9) Tagalog (Austronesian; Southeast Asia; Kroeger 1993)

a. bumili ang=lalake ng=isda sa=tindahan.

PFV.ACT.buy PROX=man OBL=fish LOC=store

'The man bought fish at the/a store.'

- b. *binili ng=lalake ang=isda sa=tindahan.*
 PFV.PAT.buy OBL=man PROX=fish LOC=store

‘The/a man bought *the* fish at the/a store.’

- c. *binilhan ng=lalake ng=isda ang=tindahan*
 PFV.DAT.buy OBL=man OBL=fish PROX=store

‘The/a man bought fish *at the* store.’

The verb indicates the role that the proximate NP plays in the clause — in our examples this is A (indicated by the ‘active’ or antipassive voice in 9a), O (‘patientive’ or passive voice in 9b) or G (‘dative/locative’ voice in 9c), but other roles are possible as well. The non-proximate NPs are either marked as oblique (*ng=*) or by the more specific case clitic *sa=* ‘locative, dative’.

What is more common is proximate GRs that only admit arguments but no adjuncts.⁷ This is found in a number of languages of the Americas (cf. Zúñiga 2006 for a survey). In Algonquian languages, for example, the most topical argument is assigned the (zero-marked) proximate GR while the other argument(s) are marked as obviative. In Central Ojibwa the proximate is furthermore referenced by raising and other constructions (see Section 4.5 below).

(10) Central Ojibwa (Algic; Eastern North America; Rhodes 1976)

- a. *aw aniniw w-gii-waabam-aa-n niw kweew-an.*
 DEM.PROX man 3-PST-DIR-3OBV DEM.OBV woman-OBV

‘*The* man saw the woman.’

- b. *aw kweew w-gii-waabam-igw-an niw aniniw-an.*
 DEM.PROX woman 3-PST-see-INV-3OBV DEM.OBV man-OBV

‘The man saw *the* woman.’

In example (10a), the A argument is assigned the proximate relation; in (10b) this is the O argument. Similarly to what we found in Tagalog, the verb morphology tracks this role assignment: the ‘direct’ suffix (*-aa*) signals that the proximate GR is the A argument (10a); the ‘inverse’ suffix (*-igw*) indicates that the proximate GR is the O argument or an inverse scenario (10b).⁸

In many languages (but not, e.g. Tagalog), hierarchical alignments are ‘frozen’ in the sense that the proximate GR choice is dictated by a hierarchy ranking speech act participants (SAP) above third persons, or possessors above possessors. (This is sometimes referred to as ‘semantic’, as opposed to ‘pragmatic’ inversion, cf. Givón 2001.) In Ojibwe, for example, SAP arguments must always be proximate while inanimate or possessed NPs must always be obviative. Thus, in order to say ‘I see him’, the first person must be proximate and the verb must be inflected as ‘direct’, indicating that the proximate is the A argument (*n-waabam-aa* [1-see-DIR] ‘I see

⁷ but under one analysis (Foley 1998), what looks like an adjunct (‘at the store’) in sentences like (9c) is in fact an argument licensed by what is traditionally called the ‘focus’ or ‘voice’ marker on the verb (here, the ‘dative/locative’ voice assigning a goal or other locational role to the proximate argument).

⁸ Marking SR-to-GR mapping under hierarchical GR choice is not the only function of inverse-marking. In some languages, it reflects deictic and empathy functions independently of the GR system (cf. DeLancey 1981; Bickel 1995; Zúñiga 2006). Conversely, hierarchical GR choice can be found without inverse-marking (DeLancey 1981; Siewierska 2004).

him’). In order to express ‘he sees me’, the first person must again be assigned the proximative GR; that it is now in O role must then be signalled by inverse inflection (*n-waabam-igw* [1-see-INV] ‘he sees me’). Inanimate arguments must always bear the obviative GR. Consider the following data:

(11) Central Ojibwa (Rhodes 1994)

- a. *w-gii-miigshkaa-go-on* *mtigo-on* *nJohn.*
 3-PST-hit.the.mark-INV-3OBV tree-OBV John.PROX
- b. **w-gii-miigshkaw-aa-n* *nJohn-an* *mtig.*
 3-PST-hit.the.mark-3OBV John.OBV tree.PROX
- ‘The tree hit John.’

In (11a), *mtigoon* ‘tree’ is in the obviative, and in order to signal that it is in O role, the verb is marked as inverse. Assigning ‘tree’ to the proximative GR and, accordingly, use a direct form, is ungrammatical, as shown by (11b). Languages differ as to whether assignment to the proximative and obviative GRs is dictated by a strict hierarchy (as for example in a number of Tibeto-Burman languages: DeLancey 1981) or whether the speaker is free to choose on pragmatic grounds of relative topicality (as for example in Tagalog), or whether both patterns coexist (as in Algonquian languages).

The data surveyed here suggest a common principle in the way referential features affect GR specifications. Regardless of whether we are looking at subjects, objects, {S,O}-relations, or proximatives, and regardless of whether we are looking at case assignment or agreement rules, it appears that many languages open their GRs preferably to animates than inanimates; to speech act participants than third persons; to known than unknown referents. These rankings can be summarized in terms of what is variously known as the ‘referential’, ‘animacy’, ‘person’ or ‘indexicality’ hierarchies (cf. among others, Silverstein 1976; Moravcsik 1978; Comrie 1981; DeLancey 1981; DuBois 1987; Givón 2001; Siewierska 2004; Haspelmath 2005; Bickel and Nichols 2007):

- (12) a. SPEECH ACT PARTICIPANT > KIN/NAME > HUMAN > ANIMATE > INANIMATE > MASS
- b. SPECIFIC > NONSPECIFIC REFERENTIAL > GENERIC/NONREFERENTIAL
- c. KNOWN/TOPICAL/THEMATIC/DEFINITE > NEW/FOCAL/RHEMATIC/INDEFINITE
- d. SINGULAR > PLURAL

But other referential notions may also play a role for GR specifications. In a number of Kiranti languages (Sino-Tibetan; Himalayas), for example, issues of politeness (face-saving) require that first person O arguments must not be overtly indexed. These languages have obligatory object agreement and the only way to delete reference to a first person O argument is to deny it object status, so that it can no longer trigger agreement. In Puma (Bickel and Gaenszle 2005), this is achieved by antipassivization: an antipassive form like *kha-en-a* [ANTIP-hear-3sPST] ‘s/he heard (someone)’ is regularly used for first person arguments, meaning ‘s/he heard us’ (and as such can even co-occur with an independent first person object pronoun, although this may be impolite).

3. Conditioning grammatical relations

The role and reference properties just surveyed define the individual arguments that can be variously included or excluded by specific GRs. But these decisions of inclusion or exclusion — often called ‘mapping’ or ‘linking’ or ‘projecting’ procedures — can, and often are, conditioned by the nature of the larger syntactic environment, specifically by properties of the entire clause, or of the predicate.

3.1 Scenario

In some languages, the assignment of an argument to a specific GR not only depends on that argument’s role and reference properties, but also on the nature of other arguments in the clause.⁹ In other words, the assignment is conditioned by the way two or three arguments interact with each other, i.e. by the scenario they define. This is illustrated here by case-marking on pronouns in Yurok.

(13) Yurok (Algic; Western North America; Robins 1958:21)

- a. *keʔl nek ki newoh-paʔ.*
 2sNOM 1sNOM FUT see-2>1s
 ‘You will see me.’
- b. *yoʔ nek-ac ki newoh-peʔn.*
 3sNOM 1s-OBJ FUT see-3s>1s
 ‘He will see me.’

The object marker *-ac* is used only when there is a third person subject in the clause. This is so in (13b) but not in (13a). In Finnish, accusative case on objects is marked only if there is a subject NP in the clause; but not, for example, in imperative constructions, which lack an overt subject NP (Comrie 1975).

In Sahaptin, it is subject rather than object marking that is sensitive to the properties of another argument (cf. Zúñiga 2006):

(14) Umatilla Sahaptin (Plateau; Western North America; Rigsby and Rude 1996)

- a. *íwínš i-tu.xnana yáamaš-na.*
 man 3sSBJ-shot mule.deer-OBJ
 ‘The man shot a mule deer.’
- b. *íwínš-nim=nam i-q’ínu-ša.*
 man-ERG=2s 3sSBJ-see-IPFV
 ‘The man sees you.’

The A-argument is assigned an ergative-marked relation only if the O-argument is a speech act participant. In (14), this condition only obtains in (b). A similar distribution is found in Tauya, a language of Papua New Guinea:

⁹ This was first identified by Silverstein 1976 in terms of ‘global rules’ of case-assignment.

(15) Tauya (Trans-New-Guinea; Papua New Guinea; MacDonald 1990)

- a. *ya-ni fanu yau-e-ʔa.*
 1s-ERG man [3sO-]see-1sA-IND

‘I saw the man.’

- b. *ya pai yau-eʔa.*
 1sNOM pig [3sO-]see-1sA-IND

‘I saw the pig.’

If there is a human or other high-ranking O-argument, the A-argument must be in the ergative, as in (15a); if not, ergative-marking is optional, as in (15b). Variations on this theme can be found in languages like Fore (also Trans-New-Guinea; Scott 1978), Acehnese (Austronesian; Durie 1987) or Rapanui (Austronesian; Du Feu 1996), where ergative marking appears whenever the O-argument precedes the A-arguments in linear order, and therefore tends to be higher ranked (in terms of the hierarchy in (12) above).

These kinds of conditions on GR-assignment are most common in case marking systems. The reason is perhaps that case (as opposed to other manifestations of GRs) often has a prominent discourse function of distinguishing transitive arguments, especially when both are animate (see Comrie 1981, among others).

3.2 Lexical predicate class

Another factor in conditioning GRs is the lexical (or lexical-semantic) class of the predicate from which arguments are selected. This is very common in the languages of the world, and there are many ways in which classes can play a role.

One way in which predicate classes define GRs is known as ‘split intransitivity’. The basic observation is that in some languages the S argument of some predicates (e.g., depending on the language, those with agentive or activity semantics) aligns with A, while the S argument of other predicates (with patientive or stative semantics) aligns with O, T or G, or a combination of these. Instead of, or in addition to such distinctions, one also often finds a class of intransitives that aligns S with G (typically with experiential semantics). Languages vary strongly as to how they group the lexicon here – indeed, they vary even as to whether the classification is rigid (‘split-S’ in Dixon’s 1994 terms), or whether it is more amenable to constructional and conceptual choice (‘fluid-S’), or whether the semantic motivation between classes draws more on notions of agentivity or experience, or on Aktionsart notions of activity (or combinations of all these). And if the classification is rigid, languages may distinguish a closed (small) vs. open (large) class (Merlan 1985).¹⁰

An example with a three-way contrast is Chickasaw. While case assignment and switch-reference (cf. Section 4.6) are based on a subject vs. object distinction, agreement is triggered by three distinct GRs: type (a) aligns S with A, type (b) aligns S

¹⁰ For recent surveys and discussion, see Donohue and Wichmann 2007. Split intransitivity is sometimes taken to challenge the universality of the notion ‘S’. But S is defined here purely by numerical valence, as an argument licensed by an intransitive predicate (cf. Section 2.1 above) and in all languages with split intransitivity that I am aware of, intransitive verbs behave differently from transitive verbs in at least some morphological or syntactic effects, minimally with regard to the number of syntactic argument positions they license. The universality of S can only be challenged by demonstrating that the difference between intransitive and transitive predicates plays no role whatsoever in a language.

with O and type (c) aligns S with G. The choice is largely lexical, but some predicates are flexible. (I use the role labels ACT ‘active’, PAT ‘patientive’, DAT ‘dative’.)

(16) Chickasaw (Muskogean; Eastern North America; Munro and Gordon 1982)

- | | |
|--|---|
| a. <i>malili-li</i>
run-1sACT
‘I ran.’ | a´. <i>chi-sso-li</i>
2sPAT-hit-1sACT
‘I hit you.’ |
| b. <i>sa-chokma</i>
1sPAT-good
‘I’m good.’ | b´. <i>is-sa-thaana</i>
2sACT-1sPAT-know
‘You know me.’ |
| c. <i>an-takho’bi</i>
1sDAT-lazy
‘I’m lazy.’ | c´. <i>iss-am-a</i>
2sACT-1sDAT-give
‘You give it to me.’ |

The data in (16) show on the left-hand side the three types of intransitive predicates; the data on the right show the transitive clauses, each highlighting the argument that is aligned with the S argument on the left.

The alignment of S with G in type (16c) reflects a frequent pattern cross-linguistically, especially for predicates that include experiential or possessive semantics, and it is often based on a metaphorical analogy of experiencers with goals (Bickel 2004b, Nichols 2007). An example is Nepali, where some experiential predicates include their S argument in a nominative-marked {O-low,T}-relation while others include it in the dative-marked {O-high,G}-relation. (To highlight the parallel to Chickasaw, I chose here similar predicates.)

(17) Nepali

- | |
|--|
| a. <i>ma</i> <i>rāmro</i> <i>thiẽ.</i>
1sNOM good COP1sPST
‘I was good.’ |
| b. <i>malāi</i> <i>alchi</i> <i>lāgyo</i>
1sDAT lazy be3sPST
‘I was lazy.’ |

In addition, there is a small set of intransitive predicates denoting body functions (‘cough’, ‘urinate’, ‘vomit’, etc.) that require ergative case, yielding, for this (and only this) predicate class, an ergative-marked {S_a, A-low}-relation (where the subscript *a* indexes the lexical class):

- (18) *kānchā-le* *khok-yo.*
last.born-ERG cough-3sPST
‘Kancha coughed.’

Predicate classification is not limited to intransitive predicates. GR definitions are affected by such classifications just as easily in transitive predicates. For one thing, transitives often treat the A argument of experiential predicates in a distinct way and assign them the same dative-marked GR as with experiential intransitive predicates.

This can again be illustrated by Nepali. Apart from sentences like (17b), there are transitives like (19):

- (19) *malāi tyo ciyā dherai man par-yo.*
 1sDAT DEM tea.NOM very please-3sPST
 ‘I liked that tea very much.’

The result of this is a complete reversal of relations in case assignment: the A argument (the more actor-like experiencer) is coded like an {O-high, G}-object, whereas the O argument (the less actor-like stimulus) is treated like a {S, O-low, T}-subject for the purposes of nominative case-assignment. (For the purpose of verb agreement, the O argument qualifies as a {S,A}-relation, as evidenced by the third person singular agreement in (19)).

Another way in which lexical distinctions among transitives matter for GR definitions is the way O arguments align with the G or T of ditransitives. In some European languages, we find lexical contrasts between {O,T}-transitives assigning accusative (e.g. German *unterstützen* ‘support’) and {O,G}-transitives assigning dative to their O argument (e.g. *helfen* ‘help’).

Other important lexical classes are motion and especially caused motion verbs. Motion verbs often assign their goal argument variably to a regular object relation or to an oblique function. English has fluid objecthood here, cf. , e.g. *load hay onto the truck*, where the goal is an oblique, with *load the truck with hay*, where the goal is treated as a direct object. By contrast, some languages fairly consistently assign their goal arguments to objects, stranding the T argument as an oblique. (For a survey see Bickel and Nichols in press).

3.3 Tense, aspect, and other clause type categories

In a number of languages, the choice between different sets of GRs, especially GRs in case assignment, depends on the choice of the verb form. We already noted this in the data from Nepali, which are fairly typical for the way such choices are distributed. Table 2 summarizes the facts (based on Clark 1989; Pokharel 2054, and my own observations), where ‘low’ and ‘high’ mean probabilistic values on the hierarchies in (12) and the subscripts *e* and *a* index arguments of specific lexically-defined predicate classes (*e* for broadly ‘experiential’ and *a* for a subset of body-function predicates):

	<i>Set I forms</i>	<i>Set II forms</i>
NOM (∅)	{S, O-low, O _e , T}	{S, A-high, O-low, O _e , T}
ERG (-le)	{S _a , A}	{S _a , A-low}
DAT (-lāi)	{S _e , A _e , O-high, G}	

Table 2: Distribution of GRs in Nepali case assignment rules

Set I includes the past (perfective or imperfective), perfect and converb forms, as well as infinitival clauses (though dialects vary in this last regard). Set II includes all other forms.

The reasons for distributions like these are best found in their etymology. In many Indo-Aryan languages, Nepali among them, one important observation is that the Set I forms go back to periphrastic participial constructions of the kind ‘with-me

the-book is written’ = ‘I have the book written’, where the agent was coded by an oblique case marker. This has developed over time into regular perfects and further into plain past tense forms (cf. ‘I have the book written’ > ‘I have written the book’) (see Peterson 1998).

In other languages, aspectual conditions are not mediated by periphrasis. In Yukatek Maya, for example, aspectual choice conditions the GRs for verb agreement in the following way: agreement follows an {S,O} pattern in clauses with perfective forms, but an {S,A} pattern in clauses with imperfective forms. In addition, the {S,O}-alignment is also conditioned by subjunctive forms, which characterize subordinate clauses.

Thus, apart from aspectual and temporal conditions, the status of clauses as subordinate vs. main, may also be a relevant factor. In other Mayan languages, e.g. Mam (England 1983) this is the only factor. Languages differ strongly in the precise definition of these conditions, and the way they interact with each other.

Another frequent way in which the clause type is relevant for GR assignment is finiteness. In many languages, nonfinite constructions obligatorily demote S and A arguments to GRs with oblique case marking. One instance of this is participial constructions in classical Indo-European languages, where overt S or A arguments must appear in an oblique case. Ancient Greek chose the genitive (while Latin chose the ablative). This is exemplified in the following (20) by the pronoun *autoû* ‘of him’, which is the S argument of participial form *asthenēsantos* ‘being feeble’:

(20) Ancient Greek (Bickel 1999)

[<i>asthenēsa-nt-os aut-oû</i>]	<i>oudépote</i>	<i>ap-é-leip-e</i>
feeble-IP-GENsM 3-GENsM	never	away-PST-leave-3sIPFV

<i>tòn</i>	<i>pápp-on.</i> (Xen. Cyr. I, 4, 2)
ART.ACCsM	grandfather-ACCs

‘When he (grandfather) was sick, he would never leave his grandfather.’

Some modern Indo-Aryan languages require overt S or A arguments in participial or converbial clauses to be in the genitive or dative, while the usual nominative or ergative case assignments are banned. Maithili chooses the dative:

(21) Maithili (Bickel and Yādava 2000)

a. <i>Rām-kē</i> / * <i>Rām</i>	<i>ehan kitāb</i>	<i>paḍh-ab</i>	<i>ṭhīk</i>	<i>nahi</i>	<i>ai-ch.</i>
R.-DAT R.NOM	such book.NOM	read-INF	right	not	3-be

‘It is not good for Ram to read such a book.’

b. <i>hamrā</i> / * <i>ham</i>	<i>ghar āib-kē</i>	<i>pitā-jī</i>	<i>khuṣī</i>	<i>he-t-āh.</i>
1DAT 1NOM	home come-CONV	father-HON.NOM	happy	be(come)-FUT-3HON.NOM

‘When I come home, father will be happy.’

Thus, in Maithili nonfinite constructions (infinitives in *-ab* as in 21a or converbs in *-kē* as in 21b), case assignment rules align S and A with G, neutralizing distinctions made in finite constructions.

4. GR-constructions: a survey

In the preceding sections we have surveyed various ways in which languages subset arguments into GRs, illustrated exclusively by case assignment and agreement rules. But these are of course not the only kind of constructions that specify a GR. In principle any syntactic construction can specify a GR: whenever some combinatorial rule or constraint is limited to a subset of arguments, this reflects a GR. In the following, I review the best studied of these GR-constructions beyond case and agreement.

4.1 Phrase structure

Phrase structure has been noted to reference GRs in two basic ways. One of them is well-known because it is found in English. Here, clause-level phrase-structure specifies rigid positions for subject and object relations, e.g. a preverbal position for subjects and a postverbal position for objects. If GRs have such positional properties in a language, the language is sometimes said to be ‘configurational’ (following a tradition established by Hale 1983 and standard in most theories), but note that in those theories that seek to represent all dimensions of syntax in phrase structure terms, the term ‘configurational’ is also used in a different sense, as implying that there is a subject/object asymmetry at all, regardless of surface positioning possibilities (e.g. Speas 1990 or Baker 1996).¹¹

An additional property of the GR positions in a language like English is that they need to be filled obligatorily (unless they are deleted in specific constructions; see below). Obligatorily filled positions are not very common in the languages of the world, and when they occur, they need not be defined as subjects or objects. In Movima, a language isolate from Bolivia, for example, clauses have to have one overt NP (lexical or pronominal), and this is the proximative NP, i.e. the argument that ranks highest on the referential hierarchy (Haude 2006). The role of this argument as A or O in transitive clauses is then indicated by direct vs. inverse voice on the verb, similarly to what we saw in Ojibwa in Section 2.2. The same basic principle, with an obligatory proximative NP, but with a more extensive voice system, is known from Tagalog. In the example in (9), the *ang*-NP is obligatory.

Another, less well-known way in which phrase structure can reflect GRs concerns projection levels, in particular the NP vs. N distinction. Belhare (Bickel 2004a, 2006), an Eastern Kiranti language of Nepal, for example, has two kinds of primary object-GRs: one for specific {O,G} arguments, and one for generic {O,G} arguments. Specific objects project a full-fledged NP; generic objects only bare Ns (for a similar pattern in related Limbu, see Angdembe 1998):

(22) Belhare (Sino-Tibetan; Himalayas; Bickel 2004a, 2006)

- a. *unchik-ŋa* [_{NP} *khaĩ=kha cece*] *n-cai-t-u*.
 3ns-ERG good=ART meat 3nsA-eat-NPST-3sO

‘They eat (the) good meat.’ [specific referent]

- b. *unchik* [_N (**khaĩ=kha*) *cece*] *n-ca-yu*.
 3nsNOM good=ART meat 3nsS-eat-NPST

‘They eat meat.’ (= ‘They are not vegetarians.’)

¹¹ In such theories, free ordering of GR-bearing NPs (i.e. apparent nonconfigurationality) is usually accounted for by constraints against NPs in argument positions, so that the freely ordered NPs are no longer real arguments.

In (22a), the O–argument is specific and is therefore realized as a specific object relation. As such, it can be expanded into a modified NP. In (22b), by contrast, the O–argument is generic and is therefore realized as a generic object; as such it cannot be expanded into a modified NP. The same distinction between specific and generic objects is also relevant for agreement: only specific objects trigger agreement, as shown by the different verb forms in (22). However, the distinction is irrelevant for all other GR-constructions of the language: all primary objects are assigned absolutive case, they can be fronted (e.g. instead of *unchik cece ncayu*, one can also say *cece unchik ncayu*), and they can be relativized on (see Bickel 2006).

4.2 Diathesis

Many kinds of diathesis (voice, applicatives, causatives, etc.) assign virtually any argument to some specific derived GR. Many applicatives, for example, are able to assign virtually any argument or even adjunct to object status (German *be-* for example can turn a locative adjunct into a direct object just as well as a G argument, cf. the applicativized locative adjunct in *be-arbeiten* ‘work on or at sth.’ from *arbeiten* ‘work’ with the applicativized G argument in *be-schenken* ‘give so. a gift.’ from *schchenken* ‘give a gift’).

But sometimes diatheses can only assign members of one GR to another GR, and then these construction specify an ‘input’ GR. Especially passives and antipassives are often restricted in such ways. German, for example, has one passive (using the auxiliary *werden*) on arguments projectable into the {O,T} relation, and one passive (using *kriegen* or *bekommen*) on arguments projectable into the {G}-relation:

(23) German

- a. *Der Wagen wurde ihm geschenkt.*
 ART.sM.NOM car PASS.AUX.3sPST 3sDAT give.as.present.PST.PTCP
 ‘The car was given to him as a gift / for free.’
- b. *Er kriegte den Wagen geschenkt.*
 3sNOM PASS.AUX.3sPST ART.sM.ACC car give.as.present.PST.PTCP
 ‘He was given the car for free / as a gift.’

The result of passivization and antipassivization is a new set of derived and demoted argument roles. Derived S and A arguments share syntactic transitivity with their non-derived counterparts — derived-S occurs in intransitive, derived-A in monotransitive clauses —, but they differ in argument structure: derived clauses still contain two-place or three-place predicates, with A, O, T, and G roles, whereas non-derived S and A clauses contain one-place and two-place predicates, respectively. Despite this difference, most languages treat derived S and A roles exactly like non-derived S and A roles for many purposes (e.g. the roles are assigned the same case, trigger the same kind of agreement etc.). Yet, as we will see in Section 5, some constructions in some languages treat derived and non-derived S roles differently. Demoted A (as in passives) or demoted O (as in antipassives) arguments are sometimes called ‘logical subjects’ and ‘logical objects’, respectively. They typically behave like adjuncts, but for specific constructions, they can also align with other grammatical relations. Again, examples of this are discussed in Section 5.

4.3 Conjunction reduction

In many languages, such as Chinese, sentences like the following receive a natural interpretation based entirely on world knowledge (Comrie 1988):

(24) Mandarin Chinese (Sino-Tibetan; China; LaPolla 1993)

nei ge ren ba xigua diao zai dishang, sui le.
that CLF person OBJ watermelon drop LOC ground break PFV

‘That man dropped the watermelon on the ground and it burst.’

In the English translation, the sentence only receives a natural interpretation if we include the pronoun *it* in the second clause (as is done in the translation of (24)). Without *it*, the syntax of English enforces an interpretation whereby the S argument of *burst* is the same as the A argument of *drop* — despite all our world knowledge that makes this a very unlikely scenario. The reason for this is that English, but not Chinese, has a GR-construction here. The construction is conventionally called ‘conjunction reduction’. It is formally identified by deletion of the subject argument in the second clause and by a rigid constraint demanding coreference between the two subjects. It is important to note that the coreference condition is a rigid syntactic constraint on interpretation, which can even overrule pragmatic background assumptions, because conjunction reduction is easily confused with zero anaphora, which does not impose any such constraint. Zero anaphora is the wide-spread tendency across languages to leave out topical arguments, such as was done in the second clause of the Chinese version of (24). Unlike under conjunction reduction, the interpretation of zero anaphora entirely rests on our knowledge of the world and the previous discourse.¹²

Conjunction reduction is probably not very common in the languages of the world. An interesting example comes from Dyirbal, however, where the construction demands coreference of the {S,O}-arguments, i.e. reversing the English alignment:

(25) Dyirbal (Pama-Nyungan; Northern Australia; Dixon 1972)

bayi yara bangun dyugumbirubalga-n badyi-nyu.
DET.sM.NOM man.NOM DET.sF.ERG woman-ERG hit-NFUT fall.down-NFUT

‘The woman hit the man and he (*she) fell down.’

The construction is formally characterized not by a conjunction or affix but by forming a single intonation group. This distinguishes the construction from syntactically unconstrained zero-anaphora (cf. Dixon 1979b for further discussion of this important point).

In Tagalog, the deleted argument in conjunction reduction precedes the antecedent,¹³ and both arguments must bear the proximative GR, regardless of their semantic role:

¹² Conjunction reduction has been claimed, for example, for many Indo-Aryan languages. On closer inspection, however, all putative instances turn out to be zero anaphora, where world knowledge can easily override the syntax, like in Chinese and unlike in English. See Bickel and Yāda 2000 and Bickel 2004b for discussion.

¹³ When it does not, this is zero anaphora, and then there is no GR specified at all. See Kroeger 1993

(26) Tagalog (Kroeger 1993)

- a. *tinukso ng=mga=kaibiganat kinagalitan si=Juan ng=kaniya=ng guro.*
 PFV.PAT.tease OBL=p=friend and PFV.G.anger PROX=J. OBL=3sDAT=LNK teacher

‘His friends teased and his teacher scolded Juan.’

- b. *pumunta sa=tindahan at bumili ang=kapatid ko ng=bigas.*
 PFV.ACT.go LOC=store and PFV.A.buy PROX=sibling my OBL=rice

‘My brother went to the store and bought some rice.’

- c. *?/* niluto ang=pagkain at hinugasan ang=mga=pinggan ni=Josie*
 PFV.PAT.cook PROX=food and PFV.G.wash PROX=p=dish OBL=J.

Intended: ‘Josie cooked the food and washed the dishes.’

In (26a), *Juan* is chosen as proximate in the two conjoined clauses, and the verbal voice inflections (O and G) signal Juan’s role as O (patient) and G (experiencer) argument, respectively. In (26b), *kapatid ko* ‘my sibling’ bears again the proximate GR, and here the verbs indicate a role as A in each clause. (26c) is ill-formed because the proximate arguments *pagkain* ‘food’ and *mga=pinggan* ‘dishes’ is not shared and not deleted. The fact that the clauses share an agent is irrelevant.

4.4 Nonfinite constructions

The key property of conjunction reduction is that a missing argument is obligatorily interpreted as coreferent with the preceding subject (or, in Dyirbal, the {S,O}-relation), but there need not be a missing argument to begin with. It is perfectly fine not to omit any argument, regardless of whether there is coreference or not (cf. *My friend_i went to town and he_i bought a case of champagne!*). This is very different from cases where a language bans the occurrence of any overt argument in some construction. The most common such constructions involve nonfinite forms (infinitives, participles, converbs, purposives, supines, etc.), and the ban is most often specified as a ban on subject arguments. Most European languages, for example, ban the appearance of any overt subject in infinitives (cf. **he to work*) or converbs (**while he working*) — but it is important to note that this is by no means universally so: many languages allow any overt argument in, for example, infinitival clauses (e.g. Nepali: Bickel and Yādava 2000), or they allow them if they are mapped into a specific case relation (cf. Section 3.3 above).

Most bans on overt arguments are complemented by some constraint or formal marking regulating the reference of these arguments. One type involves superordinate constructions, such as control and raising construction. Another type involves morphological coreference marking. Both types also occur without obligatory argument deletion, and they are discussed in the following.

4.5 Control, raising and other coreference constructions

Many languages have constructions that require a certain subordinate GR (the ‘controllee’) to be coreferential with a superordinate GR (the ‘controller’). These constructions vary typologically in two basic ways. One variable is whether the superordinate GR is a semantic argument of the superordinate clause. If it is, the construction is traditionally called a ‘control construction’ (e.g. *he wants to go*); if not, it is called a ‘raising construction’ (*he seems to work*). Another variable concerns the

question whether the subordinate argument is obligatorily deleted (cf. the preceding section) or not. When it is obligatorily deleted, the construction is sometimes said to involve ‘EQUI-deletion’. When it is not deleted, coreference constructions are sometimes called ‘copying constructions’, ‘backward control’, ‘backward raising’, or, when combined with verb agreement in the main clause, ‘long-distance agreement’. (We will encounter examples of backward control and raising below.)

In any of these constructions, the controller is sometimes lexically defined as subject (control: *I want to work*, raising: *I seem to work*), sometimes as object (control: *I ask you to work*, raising: *I believe you to work*). The controllee, by contrast, is most often defined as subject. But other GRs are also known, especially in constructions that do not ban the occurrence of overt NPs in the subordinate clause.

In Belhare raising and control constructions with *nus-* ‘may’ and *khes-* ‘must’,¹⁴ for example, the controllee must bear the {S,O}-relation. (The controller is always S with these verbs.)

(27) Belhare (Bickel 2004a)

- a. *khon-ma nui-ka*
play-INF may.NPST-2

‘You may play.’

- b. *lu-ma nui-ka*
tell-INF may.NPST-2s

‘(They/someone) may tell you.’, not *‘You may tell someone.’

The dependent infinitive can have overt arguments (e.g. *unchikɲa* ‘they.ERG’ and *han* ‘you.NOM’, as in *unchikɲa han luma nuika* ‘they may tell you’), i.e. there is no syntactic ban on overt NPs in infinitives in this language (although Belhare speakers in general tend not to use overt NPs unless they are really unavoidable pragmatically). But regardless of whether arguments are overt or not, the constraint holds that the lower {S,O} argument must be coreferential with the S argument of the main clause: **hanna luma nuika*, with the ergative-marked pronoun *hanna* ‘you.ERG’, is ungrammatical because it would require A=S coreference: ‘you [S] may [A] tell them’. Similar patterns of syntactic ergativity have been noted in a number of Nakh-Dagestanian (Caucasus; Bickel and Nichols 2001).

Another example of an {S,O} constraint in a control construction is found in Dyirbal:

(28) Dyirbal (Dixon 1995)

- a. *bayi yara walɲgarra-nyu baɲɲun yibi-ɲgu bura-li*
DET.sM.NOM man.NOM want-NFUT DET.sF.ERG woman-ERG see-PURP

‘The man wanted the woman to see him’ (e.g. while he was ‘showing off’).

- b. *bayi yara walɲgarra-nyu bural-ɲa-ygu baɲɲun yibi-gu*
DET.sM.NOM man.NOM want-NFUT see-ANTIP-PURP DET.sF.DAT woman-DAT

‘The man wanted to see the woman’ (he might be worried about her).

¹⁴ *Nus-* ‘may’ is a control verb: it does not have an impersonal alternate and assigns a semantic role to its S argument. *Khes-* ‘must’ and some other verbs do have impersonal alternates and are likely to be raising verbs. See Bickel 2004a for discussion.

The verb *wahlgarra-* ‘want’ is an intransitive control verb and requires the lower {S,O}-argument to be coreferential with its S argument. In (28a), this is *yara* ‘man’ and this argument is in O function in the dependent clause. If the semantics require coreference with an A argument, as in (28b), the dependent clause needs to be antipassivized so that the A argument is re-assigned the S function.¹⁵

Occasionally, control or raising verbs in some languages constrain the controllee to bear a more narrowly defined GR. In a number of Mayan languages, control construction impose obligatory deletion of the controllee, and the controllee is restricted to S arguments:

(29) Yucatec (Mayan; Mexico; Verhoeven 2005)

- a. *in=k'áat bin Cancun.*
 1sA-wish go C.
 ‘I want to go to Cancun.’
- b. *in=k'áat in=kan Màaya.*
 1sA-wish 1sA-learn[-3sO] M.
 ‘I want to learn Maya.’
- c. *in=k'áat káa u=bis-en Cancun in=tàatah.*
 1sA-wish COMP 3sA-carry-1sO C. 1sPOSS=father
 ‘I want my father to bring me to Cancun.’

If the dependent clause is intransitive, as in (29a), it is integrated into a control construction: the verb is nonfinite and the controllee is obligatorily deleted under coreference (so that *in=bin* ‘I go’ would be ungrammatical here). Under all other conditions, the dependent clause obligatorily retains clitics for both arguments. Neither coreference of the subordinate A as in (29b) nor coreference of the subordinate O as in (29c) allow the use of the Yucatec control construction.

In the data so far, the controllee was specified as {S}, {S,A}, or {S,O}, but purely referential GR specifications (as discussed in Section 2.2) are also found in coreference constructions. The Algonquian language Ojibwe, for example, imposes a coreference constraint on ‘know’ constructions between the main clause obviative and the embedded proximative argument. The proximative argument is the one that is considered most topical in a clause, and it is the A argument if the verb is inflected as direct and the O argument if the verb is inflected as inverse (cf. the data in (10) above). (Note that the controllee is not deleted in this language.)

¹⁵ Dixon 1995 and Manning 1996 claim that these are not control constructions comparable to English ‘want’ constructions because the dependent clause is not embedded and because the subordinate controllee may be overt (as long as it is coreferential with the superordinate S argument). But English infinitives after *want* are not embedded either (they do not fill the canonical object position, nor have they all object properties, cf. Van Valin and LaPolla 1997: 461ff), and overt controllees are widely attested in other languages (cf. the Belhare example above and Polinsky and Potsdam 2006 for a recent survey). What makes Dyirbal ‘want’ sentences control constructions is that the matrix verb “carries the expectation of a further verb in purposive construction with it” (Dixon 1995: 206). Also, I would be surprised if the dependent clause did not exhibit such properties of subordination as disjunct illocutionary scope (whereby only one but never both clauses can for example be questioned). But on this we lack data.

(30) Central Ojibwa (Rhodes 1994)

- a. *n-gikenm-aa-g ninw-ag gii-baashkzw-aa-waad Maagiiyan.*
 1-know-DIR-3p man-pPROX PST-shoot-DIR-3 Marge-OBV

‘I know the men shot Marge.’

- b. *n-gikenim-aa Maagii gii-baashkzo-go-d ninw-an.*
 1-know-DIR[3s] Marge.PROX PST-shoot-INV-3 man-pOBV

‘I know the men shot Marge.’

- c. **n-gikenm-aa-g ninw-an gii-baashkzo-go-d Maagii.*
 1-know-DIR-3p man-pOBV PST-shoot-INV-3 Marge.PROX

Intended: ‘I know the men shot Marge.’

In (30a), *ninwag* ‘the men’ is chosen as the proximative GR and this argument is the controllee, as shown by its coreference with the third plural controller in the main clause (indexed by the agreement suffix *-g* ‘3p’ and registered as an obviative O-argument in the main clause by the direct marker). In (30b), *Magi* ‘Marge’ is chosen as the proximative GR in the dependent clause. Accordingly, this argument is now the controllee, and as such, is coreferential with the third person singular obviative controller in the main clause. The construction is ungrammatical, however, if the coreference relation is intended as holding between the main clause controller and the subordinate obviative argument. This would be the case in (30c), where the main clause inflection signals a plural controller but in the subordinate clause it is again the singular NP *Maagii* that is assigned the proximative GR. A similar pattern is found in Tagalog, where the controllee in raising constructions must also bear the proximative GR. Different from Ojibwe, the Tagalog raising construction also requires the controller to be in proximative function (and also different from Ojibwe, it must be deleted in the dependent clause; see Kroeger 1993).

4.6 Switch-reference and other kinds of cross-clausal coreference marking

Many languages have a morphological device for explicitly signalling whether or not selected arguments of two clauses have the same reference or not. Such devices are called switch-reference markers. The question of which arguments are monitored for coreference is defined by the GR of the construction. The near-universal choice here is subject relations, i.e. switch-reference morphology indicates coreference of subjects. The following illustrates this in Kâte:

(31) Kâte (Trans-New-Guinea, Papua New Guinea; Pilhofer 1933)

- ra fisi-pie fahare-râ yâpe?-yopa-pie mafa-yenî?*
 go arrive-SEQ.3pDS rise-SEQ.SS chase.away-3pDO-SEQ.3pDS stuff-3pPOSS

- behe-râ wise-pie fiu? ro=fâre-mbiŋ.*
 throw.away-SEQ.SS flee-SEQ.3pDS illicitly take=all-3pREMOTE.PAST

‘When they_i (the foreigners) arrived, they_j (the villagers) got up and chased them away. They_i threw away their stuff and fled. Then, they_j stole their stuff.’

The dependent forms marked as ‘DS’ signal that the following clause has a different subject; those marked as ‘SS’ indicate that the same subject referent will follow.

Switch-reference systems compare the reference of two GRs, the one in the clause bearing the switch-reference markers with the one in the clause referred to by the markers. Most often the GRs are specified in the same way, both as subjects. But sometimes they need to be distinguished as the controller GR and the controllee GR, and then the controller may precede or follow the controllee, and there may be a constraint on whether one or both or none of them need to be deleted.¹⁶ Also, note that switch-reference systems may be equipollent as in Kâte, with one marker for ‘same GR’ and one for ‘different GR’, but privative systems with a marker for just one option (typically for ‘same subject’) are also very widely attested (e.g. in the form of many converbs, such as Turkish forms in *-İp* and *-ErEk*). And finally, it is important to note that switch-reference can be found in many different kinds of clause linkage. Kâte illustrated switch-reference in clause chaining, but switch-reference devices are also frequently found in various kinds of subordination and embedding.

Switch-reference systems in clause chaining seem to favor subjects as the GR they target, but other options are also attested. Dyirbal has a privative system marking coreference, where the controller is defined as {S,O} and the controllee as {A}. The controllee must precede the controller and the two events must follow each other immediately:

(32) Dyirbal (Dixon 1994)

yabu *ɲuma-ɲgu* *bura-n* (*ɲuma*) *banaga-ɲurra*.
mother.NOM father-ERG see-NFUT father.NOM return-{S,O}={A}

‘Father saw mother and immediately returned.’

Another, very rare kind of GR referenced by coreference marking is reported (in a brief analysis) from Angaataha, a Papuan language, which apparently has a system of switch-reference targeting locative relations rather subjects (Trans-New-Guinea; Huisman 1973; Foley 1986).

Outside chaining constructions, the GRs monitored by switch-reference are more varied. Eskimo languages, for example, have two forms of signalling coreference with a subject controller: one for subject controllees and one for object controllees (traditionally called ‘reflexives’):

(33) Central Yup’ik Eskimo (Eskimo-Aleut; Alaska; Reed 1977)

a. *angute-m* *tange-llr-ani* *tuntuvak* *aya-llr-uuq*.
man-ERG see-WHEN-3sA>3sO.{O}={S,A} moose.NOM go.away-PST-INTR-3s

‘When the man saw him_i, the moose_i went away.’

b. *tang-llr-miniu* *tuntuvak* *angun* *aya-llr-uuq*.
see-WHEN-3sA>3sO.{A}={S,A} moose.NOM man.NOM go.away-PST-INTR.3s

‘When he_i saw the moose, the man_i went away.’

The mirror-image of this is conjunct participles and related constructions, where the controllee is always the subject (and, in addition, obligatorily deleted) but where different forms indicate the GR of the controller in the main clause. Warlpiri has two

¹⁶ The special case of constructions with obligatorily deleted controllees and coreference-marking is sometimes identified as ‘depictive’ or ‘secondary’ predication. In Tagalog, for example, secondary predicates must have a controller bearing the proximative GR (Kroeger 1993:30f). See Schultze-Berndt and Himmelmann 2004 for a typological survey of depictive predicates.

options for signalling coreference (plus one for disjoint reference). The suffix *-karra*, illustrated by (34a), indicates that the controllee is coreferential with the subject of the main clause, while *-kurra*, as in (34b), indicates that the controllee is coreferential with the (primary) object.

(34) Warlpiri (Pama-Nyungan; Australia; Simpson 1991)

- a. *ngarrka=ka wangka-mi karli jarnti-rninja-karra.*
 man.NOM=PRS speak-NPST boomerang.NOM trim-INF-SIM.{S,A}={S,A}
 'The man talked when trimming the boomerang.'
- b. *ngajulu-rlu-rna yankirri pantu-rnu, ngapa nga-rninja-kurra.*
 1s-ERG=1sA emu.NOM spear-PST water.NOM drink-INF-{S,A}={O,G}
 'I speared the emu while it (not I) was drinking water.'

The classical Indo-European languages have as many options as they have distinct cases: the case on conjunct participle indicates with which argument or adjunct of the main clause the (obligatorily deleted) subject of the participle is coreferential: namely with the one that bears the same case.¹⁷

In canonical instances of switch-reference, the system is marked on the verb or on conjunctions. When coreference is marked on pronouns, the system is usually not called switch-reference but cross-clausal or long-distance reflexivization or logophoricity. (The term 'logophoricity' is usually reserved for clause linkage involving reported speech or thought: Hagège 1974.) Since the controlled pronoun can typically assume any GR, such systems only need to specify the GR of the controller (also known as the antecedent). Most often this is the subject, but logophoric pronouns sometimes specify their controller as whichever argument represents the information source.

A construction related to switch-reference is odd-pivot marking, described for the Australian language Kayardild by Evans (1995). Odd-pivot marking involves the spread of additional cases on subordinate clauses (and their NPs) and signals that two clauses do not share a subject referent, i.e. that they share no argument at all, or that they share one or more arguments but at least one of them is not a subject in both clauses. Thus, the relevant GR is again the subject relation.

4.7 Relativization

One type of construction that varies strongly in terms of GR specifications across languages is the relative construction. Relative constructions turn a propositional expression into a referential one, e.g. a clause like *he read it* into *the one he read*. The referent of the expression is thereby chosen among the arguments and adjuncts of the clause, and I refer to it as the relativization site. The site is sometimes linked to an NP that further constrains the referent (e.g. *the book he read*) and this NP may occur inside the construction (and then it is called 'internally headed', as in *whichever book he read*) or adjacent to it (and then it is an attributive construction, as in *the book which he read*). In many languages, relative constructions have no GR restrictions: the same construction can be used on any relativization site. But we also frequently find

¹⁷ This system of coreference-marking is complemented by what is called 'absolute' constructions. These constructions often (but not obligatorily) have an overt subject with disjoint reference. See Bickel 1999 for discussion of this point and a short typological survey of absolute constructions.

constructions that are limited to sites bearing a specific GR. One very common instance of this are attributive participles, e.g. English *the man [walking down the street]* or *the man [telling the stories]*. Often such constructions are restricted to relativization on subjects (cf. **the stories telling the man* which is ungrammatical when intended as ‘the stories that the man is telling’). However, it has been noted for languages with syntactically unconstrained site choice, that in discourse the most frequent sites tend to be O or S arguments (Fox 1987). It does not come as a surprise therefore that relative constructions are not infrequently specialized for {S,O}-GRs. Here is one example:

(35) Oirata (Timor-Alor-Pantar; Eastern Indonesia; Donohue and Brown 1999)

- a. *inte [ihar [mara-n]] asi.*
 1pNOM dog go-REL see
 ‘We saw the dog that had left.’
- b. *[ihar [ante asi-n]] mara.*
 dog 1sNOM see-REL go
 ‘The dog that I saw left.’
- c. **[ihar [ani asi-n]] mara.*
 dog 1sACC see-REL go
 ‘The dog that saw me left.’

Relativization is marked in this language by the suffix *-n* and is only possible if the site is S, as in (35a), or O, as in (35b). It is not possible to relativize on A. In order to express the intended meaning of (35c), a circumlocution is used that does not involve a relative construction but clause chaining with same-subject morphology (‘the dog saw me and left.’)

As we have seen in other constructions (e.g. raising or case assignment), referential properties can also be relevant for GR definitions. This is also attested for relative constructions. In Tagalog, the relativization site must bear the proximative GR (Foley and Van Valin 1984; Kroeger 1993). In Movima, we find the opposite. In this language, the relativization site is limited to obviatives in transitives and the S argument of intransitives. Obviatives are assigned to whichever argument is less topical in discourse, which mostly means that it ranks low on the hierarchies in (12a) and (12d) above.

(36) Movima (isolate; Bolivia; Haude 2006)

- a. *[kinos ney ay’ku [di’ jaynta kayni]]*
 ART.F.ABSENT here aunt REL DISCNT die
 ‘that aunt of mine who has already died’
- b. *[isos wa:ka [di’ chik<a>ye=is neyru=s beñ’i]]*
 ART.pPST cow REL find<DIR>=PL.ABSENT here=DET grassland
 ‘the cows which they had found in this grassland’
- c. *[us ney juyeni [di’ alwani-kaya=y’i]]*
 ART.M here person REL talk-INV=1p
 ‘that person who had spoken to us’
- d. *[is juyeni [di’ jayna kwey way-na n-i’ne]]*
 ARTp person RELDISCNT ANTIP take.up-DIR OBL-3sF
 ‘the people who had taken her up’

(36a) shows relativization on an S argument (the argument of *kayni* ‘die’). In (36b) and (36c) we find relativization on the obviative argument. The obviative is assigned the O-role (of *chikaye* ‘find’) by the direct verb inflection in (36b) and the A-role (of *alwanikaya* ‘talk to’) by the inverse verb inflection in (36c). An alternative to the strategy in (36c) is to antipassivize the verb so that the A argument (of *wayna* ‘take up’) is re-assigned derived S status and the O argument is demoted to an oblique NP. This can be observed in (36d), where the relativization site is the derived S argument of the antipassivized clause and the O argument is marked oblique (*ni’ne* ‘to her’). Relativization on A arguments is impossible if the argument is not either assigned obviative status or re-assigned to derived-S by antipassivization.

4.8 Quantifier and other floating constructions

Another construction with considerable variation in the kind of GR involved is known as ‘floating’. Floating refers to the possibility offered by some languages to position a referential operator, such as a quantifier, a numeral, or an indefinite marker away from the NP which it has scope over. The actual scope is then often regulated by a GR, i.e. the floated operator can only take NPs in its scope that bear a certain GR. In Tagalog, the quantifier *lahat* ‘all’ may float to the Wackernagel position, but then it can only have scope over the NP that bears the proximative GR:

(37) Tagalog (Schachter 1976; Kroeger 1993)

- a. *sumusalat lahat ang=mga=bata ng=mga=liham.*
 ACT.IPFV.write all PROX=p=children OBL=p=letter

‘All the children are writing letters.’

Not: *‘The children are writing all the letters.’

- b. *sumusalat lahat ng=mga=bata ang=mga=liham.*
 ACT.IPFV.write all OBL=p=children PROX=p=letter

‘The children are writing all the letters.’

Not: *‘All the children are writing letters.’

If the quantifier *lahat* is in the Wackernagel position, it does not refer to the immediately following NP but the NP bearing the proximative GR, i.e. *ang=mga=bata* ‘the children’ in (37a) and *ang=mga=liham* ‘the letters’ in (37b).

In Yéli Dnye, an isolate of Melanesia, most referential operators can float to preverbal position. If they do, their scope is defined as the NP bearing the {S,O} relation:

(38) Yéli Dnye (isolate; Rossel Island; Henderson 1995, Levinson 2003)

- a. *yeli pi nkéli=k:oo ngmê=doo=dpodo.*
 Rossel man boat=inside INDF=3REM.PST.CNT=work.CNT

‘A Rossel man was working in the boat (day before yesterday).’

- b. *pi=knî=y:oo chêêpî ngmê=dê=d:ii=ngmê.*
 man=AUG=pERG stone INDF=3PST.PUNCT=throw.PROX=MONO.S>3sO

‘People threw some stones’, not: *‘Some people threw the stones.’

The sentence in (38a) is intransitive and the floated indefinite marker *ngmê* takes scope over the S argument ('some Rossel man'). In (38b), the scope of the marker can only be over the O argument ('some stones') and not over the A argument ('some people'). Quantifier-floating restricted to {S,O} argument has also been noted for Halkomelem (Salish; Gerdts 1988), and Donohue (2007) describes it for Japanese.

4.9 Other constructions

Apart from what we surveyed so far, many languages have other constructions specifying GRs. Here are some that have been noted in the literature:

- *Lexical nominalizations*: Lexical nominalizations often remap arguments to specific GRs, producing specific kinds of alignments (e.g. *my dancing* and *my hitting*, where the possessor codes S or A, but not O).
- *Focus constructions*: In some languages, one focus construction is used for A arguments, while another one is reserved for {S,O} arguments. This is widely attested in Mayan languages (e.g. Van Valin 1981), but it has now also been reported for a language of Melanesia, Yélî Dnye (Levinson 2003).
- *Imperatives*: While in many languages, imperatives can only be formed from agentive or volitional predicates (e.g. in Tagalog, Kroeger 1993: 88ff), some languages specify the conditions in terms of a GR. In English for example, imperatives can only be formed if the addressee is the subject, regardless of its semantic role. Hence we get *watch this!* with a transitive A argument and both *go!* with a agentive S argument and *be seen at the grand opening!* with a passive derived-S argument. But we do not get imperatives where the addressee is a transitive O argument (**People see at the grand opening!* intended as 'they should see you at the grand opening') (cf. Dixon 1979a; Comrie 1981; Dixon 1994; Donohue 2007). In many languages, however, imperatives are a regular subset of agreement paradigms, and, therefore, the definition of the controller is simply the same as the definition of the agreement-triggering GR. No special statement is needed in the grammar.

Another construction type that is frequently adduced as GR-specifying is intraclausal reflexives, but the evidence is often thin for this. Because reflexive pronouns often lack a nominative form in many languages, it is commonly expected that they only take subjects as their antecedents. In many cases, however, antecedents can assume a variety of GRs, e.g. both subject (*John talked about himself*) and objects (*John told Mary about herself*) and it is not at all clear whether the relevant constraints are best captured in terms of GRs. Further, closer inspection of languages for which a subject antecedent condition has been claimed, suggest that antecedent choice is flexible and influenced by such discourse variables as topicality and animacy and by lexical choices (e.g. experiential vs. other predicates) (cf. a 2002 discussion on the Linguistic Typology mailing list (LINGTYP) about cases of reflexives in A-function, and for example Timberlake 1980; Faarlund 1998; Bickel and Yādava 2000 or Bickel 2004b for evidence against a strict definition of reflexive antecedents as subject in a number of languages). This is not to say that reflexive antecedents are never defined by a strictly syntactic notion of subjects, but the burden of proof is much heavier than is often assumed.

5. Coding constructions and GR specifications in other constructions

Expanding on a suggestion of Keenan's (1976), one can classify GR-constructions as coding vs. behavioral constructions. Under coding constructions I include here case, agreement, phrase structure and diathesis — they all have in common that they overtly mark the GR that they specify, by assigning a specific case, selecting a specific agreement paradigm, defining a specific position in phrase structure, or signalling a specific assignment of roles to a GR (diathesis). All other constructions surveyed are behavioral construction insofar as the GR they specify is only relevant as a constraint on syntactic behavior (e.g. on what can or must be deleted, on what can be relativized on etc.) but there is no overt indication of the GR in terms of morphology or position.

The distinction is important in some but not all languages. It is important whenever the GRs of behavioral constructions are constrained by simultaneously established GRs of coding constructions. Thus, if an argument is affected by diathesis, triggers a specific agreement paradigm, appears in a specific case or is assigned a specific position, this sometimes has an impact of whether or not the argument is included in a GR specified for another construction such as raising or relativization. Languages differ strongly in these regards.

For example, under detransitivizing diathesis (passive and antipassives), the derived S argument is sometimes not admitted to the same GR as non-derived S arguments. In Section 4.5, we noted that Yucatec control constructions are possible only if the controllee assumes the {S}-relation (cf. the data in (29)). S arguments that are derived by one kind of passivization, however, do not qualify and as a result cannot be deleted controllees like ordinary S arguments:

(39) Yucatec (Verhoeven 2005 and personal communication)

- a. **in=ka'áat bis-a'l Cancun tumen in=tàatah.*
 1sA=wish carry-PASS.IPFV C. by 1sPOSS=father
- b. *in=ka'áat bis-bil Cancun tumen in=tàatah.*
 1sA=wish carry-GER.PASS C. by 1sPOSS=father
- 'I want to be taken to Cancun by my father.'

Under regular passivization, as in (39a), the derived S argument of the dependent clause cannot be equi-deleted. But Yucatec has an alternative passive construction, the gerundial passive in (39b), and the derived S of this construction groups with non-derived S arguments. The GR specified by the Yucatec control construction is therefore defined as the set {non-derived-S, gerundial-passive-S} rather than as a generalizing notion {S, derived-S}. Such kinds of specifications are essentially parallel to other conditions on argument subsetting discussed in Section 2.2 above.

Constructions not only vary as to whether a derived S argument is part of an S-including GR, but also as to whether a demoted (oblique) argument (a demoted O in antipassives, a demoted A in passives) is included in a GR. In most languages, the demoted A in passives does not qualify as a subject in such constructions as switch-reference or conjunction reduction. But in some languages, it does. Consider the following data from Seri, where disjunct subject reference is signalled by the clause-final particle *ma* 'different subject' and co-reference by the absence of this particle:

(40) Seri (Hokan; Mexico; Farrell et al. 1991)

- a. *mi-naɪt̚ kom m-po-ki:xk x, ʔata:p ko-mi-si-a: ʔa=ʔa.*
 2POSS-skin the 2sSBJ-IRR-wet AUX mucus 3OBL-2sSBJ-IRR-be AUX=DECL
 ‘If you wet your skin, you will be with mucus.’ (i.e. get a cold)
- b. *ʔa:t kiʔ p-a:ʔ-ka: x, ʔe:poʔ kiʔ mos si-a:ʔ-ka: ʔa=ʔa.*
 limberbush the IRR-PASS-seek AUX ratany the also IRR-PASS-seek AUX=DECL
 ‘If limberbush is looked for, white ratany should be looked for also.’
- c. *m-yo-a:ʔ-kašni, kokašni šo m-t-aʔo ma.*
 2sSBJ-DIST-PASS-bite snake a 2sSBJ-REAL-see DS
 ‘You were bitten, after you had seen a snake.’

(40a) illustrates a regular same-subject relation holding between the A of the subordinate (conditional) and the A of the main clause. In (40b) the clauses are both passivized, but now the same-subject relation holds between the demoted (and deleted) A arguments. The fact that the derived S arguments are different (the two kinds of plant named *ʔa:t* and *ʔe:poʔ*, respectively), is immaterial. Conversely, the fact that in (40c), the derived S of the main clause is coreferential with the A argument of the subordinate clause is irrelevant for switch-reference. What matters is that the demoted A argument of the first clause is distinct from the A argument in the subsequent subordinate clause. Therefore the subordinate clause (obligatorily) receives different subject marking (by *ma* at the end of the clause). Thus, the GR targeted by switch-reference in this language is specified as {nonderived-S, A, demoted-A}.

Case assignment rules provide other coding constructions that in some languages affect the GR specifications of behavioral constructions, while in other languages they do not. In many languages, for example, the GRs defined by relative constructions are immune to variance in case assignments resulting from lexical predicate classification (in the sense discussed in Section 3.2). In Belhare transitive clauses, for example, the most actor-like argument is part of the {S,A} relation (for example in nonfinite constructions), the other argument is part of the {S,O}-relation (for example for internally headed relativization and control constructions). For these assignments, the GRs and alignments defined by case do not matter: the S and A argument are part of the {S,A} relation regardless of whether the lexical predicate assigns it nominative, ergative or genitive. This is illustrated here by active participle constructions, which limit the relativization site to {S,A}.

(41) Belhare

- a. *un ija lim-yu.*
 3sNOM beer.sNOM [3sS-]like-NPST
 ‘S/he likes (the) beer.’
- a'. *ija ka-lim-ba*
 beer.sNOM ACT.PTCP-go-M
 ‘the one who likes (the) beer.’
- b. *(un-na) tombhira kiiʔ-t-u.*
 3s-ERG lynx.sNOM [3sA-]fear-NPST-3sO
 ‘S/he fears (the) lynx.’

b'. *tombhira ka-kit-pa*
lynx.sNOM ACT.PTCP-fear-M

'the one who fears (the) lynx' (not 'the lynx that s/he fears')

c. (*un-naha*) *u-kipma ka?-yu*
3s-GEN 3sPOSS-fear [3sS-]come.up-NPST

'S/he is afraid.'

c'. *u-kipma ka-kat-pa*
3sPOSS-fear ACT.PTCP-come.up-M

'the one who is afraid'

The data in (a), (b) and (c) of (41) illustrate A and S arguments bearing nominative, ergative and genitive-marked GRs, respectively, but they all qualify equally well for the subject GR in the participle construction derived from these clauses. As a result, the GRs specified by case assignment rules are totally different from the GR defined by the participle construction and, for that matter, from the GR of any other construction in the language (a phenomenon called 'hidden syntax' in Bickel 2004a).

This contrasts with almost all Indo-European languages, where the most actor-like argument of a transitive verb can only function as a subject if it is also assigned nominative case. For example, in German, an experiencer can only function as the subject in active participle constructions if it is in the nominative:

(42) a. *Die Studenten mög-en den Wein.*
ART.pNOM student.p like-3pNPST ART.MsACC wine.s

'The students like the wine.'

a'. *die den Wein mög-end-en Studenten*
ART.pNOM ART.M.sACC wine.s like-ACT.PTCP-pNOM student.p

'the students who like the wine'

b. *Den Studenten schmeck-t der Wein.*
ART.M.pDAT student.p taste-3sNPST ART.MsNOM wine.s

'The students like the wine.'

b'. **die der Wein schmeck-end-en Studenten*
ART.pNOM ART.sNOM wine.s taste-ACT.PTCP-pNOM student.p

Intended: 'The students who like the wine.'

Like in Belhare, lexical conditions assign A arguments to one case for some predicates and to another case for others. In (42a), the verb assigns the A argument to a nominative-marked subject GR; in (42b) it assigns it to a dative-marked $\{S_e, A_e, G\}$ -relation (where the subscript *e* indexes the lexical class). The GR specification of active participles follows this, and allows an A argument to satisfy its crucial subject GR only if it also has subject status for case-marking purposes. Therefore, relativization is impossible in (42b'). As a result, the GR identified by nominative case is virtually identical with the GR specified by the participle construction. In fact, this overlap permeates almost all GR-constructions in the language (evident in the examples by the fact that verb agreement is also controlled by the nominative-marked subject, and not by the $\{S_e, A_e, G\}$ -relation). Given this, it has occasionally even been proposed to replace the GR notion of subject in such languages by the case notion 'nominative'

tout court (e.g. Reis 1982). The empirical facts are very typical for Indo-European languages in general (see Bickel 2004b for Indo-Aryan data), and it is interesting to note that Pāṇini did not use a notion of GR in his grammar of Sanskrit (Kiparsky 2002). All that he needed were generalized semantic role (the *kāraṅgas*) and the morphological exponents of case and voice (and, of course, an intricate theory of linking).

Lexical conditions affect agreement construction just as easily as case constructions, and the typological question again arises whether differences in agreement GRs can also affect the GRs in other constructions. In Chickasaw, the language illustrated in Section 3.2, they do not. Thus, while different predicates condition different agreement GRs, switch reference (and nominative case assignment) constructions all reference a subject GR completely independent of this:

(43) Chickasaw (Munro and Gordon 1982)

- a. *top-at tiwwa-li-kat sa-hotolhko-tok.*
 bed-NOM lie-1sA-SUB.SS 1sO-cough-PST
 ‘Lying in bed, I coughed.’
- b. *alhponi’ aa-sa-bashafa-kā Bonnie-akot sa-bashaffi-tok.*
 kitchen LOC-1sO-be.cut-SUB.DS B.-CONTR.NOM 1sO-cut-PST
 ‘I got cut in the kitchen, and Bonnie did it.’

The sentence in (43a) shows same-subject marking on the subordinate clause, which shows that the difference between aligning S with A in the case of *tiwwa* ‘lie’ and with O in the case of *hotolhko* ‘cough’ is immaterial to the relevant notion of {S,A} monitored by same-subject marking here. In (43b) we find different subject marking because there is no coreference between the S of the subordinate and the A of the final clause. The fact that the S of the subordinate clause happens to be coreferential with the O of the final clause and that it even happens to trigger the same agreement forms, is irrelevant. Similar facts hold for Papuan languages like Amele (Roberts 1988) or Usan (Reesink 1983).

But in other languages and other constructions, agreement GRs conditioned by lexical classes are sometimes relevant for GRs in other constructions. Acehnese, for example, has intransitive verbs aligning S with A, as in (44a), and others aligning S with O, as in (44b). (44c) shows the respective A and O agreement markers in a transitive clause, for comparison.

(44) Acehnese (Austronesian; Sumatra; Durie 1985, 1987)

- a. *ji-jak gopnyan*
 3.A-go 3HON
 ‘S/he goes.’
- b. *gopnyan rhët-geuh*
 3HON fall-3HON.O
 ‘S/he fell.’
- c. *gopnyan ka-ji-poh-geuh.*
 3HON INCH-3A-hit-3HON.O
 ‘(S/he) hit him/her.’

Exactly the same split of GRs is also referenced by control constructions:

- (45) a. *gopnyan geu-tém jak.*
 3HON 3HON.A-want go
 ‘S/he wants to go.’
- b. *geu-tém taguen bu.*
 3HON.A-want cook rice
 ‘S/he wants to cook rice.’
- c. **gopnyan geu-tém rhët.*
 3HON 3HON.A-want fall
 Intended: ‘S/he wants to fall.’
- d. **aneuk agam nyan ji-tém geu-peuréksa lé dokto.*
 child male DEM 3-want 3A-examine ERG doctor
 Intended: ‘That boy wants to be examined by the doctor.’

The controllee is not defined here as the subject, but as as the $\{S_a, A\}$ -relation, i.e. as A for transitive predicates and as S of a subset of intransitive predicates (indexed by the subscript *a*). The S argument of this subset turns out to have similarly ‘agentive’ semantic roles as A arguments and so, the GR can equally well be called ‘Agent’, implying a close affinity to semantic notions.¹⁸ Label choice notwithstanding, what is important is that the GR specified for agreement purposes is the same as the one governing the choice of the agreement paradigm.

Last, but not least, phrase-structure rules are other coding constructions that in some languages impact the way GRs work in other constructions. As noted before, most GR-constructions in German target a uniform subject (i.e. $\{S, A\}$) relation. There is also a phrase-structure rule that assigns subjects a specific default position in the clause, the ‘prefield’ position before the finite verb indicating topicality. Subjects appear in this position (cf. (46a)), unless it is filled by another expression (as in (46b)):

- (46) a. *Sie schlief heute aus.*
 3sF.NOM sleep.3sPST today out
 ‘She slept in today.’
- b. *Heute schlief sie aus.*
 today sleep.3sPST 3sF.NOM out

GR-construction like nominative case assignment or agreement ignore this positional assignment of subjects, so that *sie* ‘she’ is assigned nominative and triggers verb agreement in both (46a) and (46b). But the position is crucial for the GR specification in conjunction reduction:

¹⁸ Relations that closely mirror semantic roles are sometimes said not be GRs at all (e.g. Van Valin and LaPolla 1997). But this is like saying that an English category like past tense is not a grammatical category just because it closely mirrors a semantic notion of past time. What matters is that the category or the relation is referenced by rules of grammar (morphological rules in the case of past tense; agreement and control rules in the case of Acehnese GRs); also cf. Dryer 1997.

- (47) a. *Sie arbeitete gestern bis spät und schlief heute aus.*
 3sF.NOM work.3sPST yesterday until late and sleep.3sPST today out
- b. **Sie arbeitete gestern bis spät und heute schlief aus.*
 3sF.NOM work.3sPST yesterday until late and today sleep.3sPST out
- ‘She worked late yesterday and slept in today.’

Deletion under coreference is only possible if the subject is in the topic-indicating prefield, as is the case in (47a). If it appears anywhere else in the clause, it cannot be deleted, as in (47b). Thus, the proper characterization of the GR targeted by conjunction reduction must not only refer to the set {S,A} but also to its prefield default position.

6. Variables and distributional trends

Table 3 summarizes the variables relevant for the specification of GRs specific languages.

Role subset:	some subset of {S, A1, A2, O, T, G}
Cast subset:	various referential notions
Conditions:	scenario, lexical predicate class, clause type properties
Construction:	various GR-specifying constructions
Coding-on-behavior constraint:	present or absent, different kinds
<i>Table 3: Summary of typological variables identifying a specific GR</i>	

Obviously, this system of variables allows enormous diversity: the role variable alone allows for $2^6 - 1 = 63$ different (non-empty) subset definitions. This is multiplied by the many ways in which various referential notions (like ‘animate’, ‘topical’, ‘speech act participant’) can further constrain or indeed directly define GRs, and external conditions (scenarios, lexical classes, tense, aspect, subordination, etc.) that can condition GR definitions. And we have seen many cases of GRs that vary from construction to construction in a single language, and if there is more than one GR construction in a language, another relevant variable is whether GR specifications in coding constructions (e.g. case assignment) affect the way other GRs (e.g. in relativization) work or not.

There have been a number of attempts to estimate significant clusterings or trends in how these variables interact. Nichols (1992) investigates areal and genealogical factors as well as correlations of GR types with word order and morphological complexity. Müller-Gotama (1994) researches the semantic role range of GRs and the relation of this with constructional choices and phrase structure types. But the topics that have dominated typological research into GRs are the role of the referential hierarchy in predicting GR types and the distribution of GRs across constructions. I take up these in turn.

6.1 Referential effects on GR distributions

When discussing how GRs can be defined, at least in part, by referential notions in Section 2.2, we noted that languages frequently reserve access to their GRs (of whatever role alignment and for whatever constructional purpose) to referents ranking highest on the referential hierarchies in (12), repeated here as (48):

- (48) a. speech act participant > kin/name > human > animate > inanimate > mass
 b. specific > nonspecific referential > generic/nonreferential
 c. known/topical/thematic/definite > new/focal/rhematic/indefinite
 d. singular > plural

Alternative principles, e.g. where access to GRs is constrained by face-saving strategies, have been less commonly noted. One reason that has been advanced to explain this trend is that GRs typically reflect grammaticalized topicality assignments and higher positions in the referential hierarchies are intrinsically more likely to be topical (Givón 2001).

The hierarchies in (48) have also been suggested to produce a specific effect on the typological distribution of case assignment. From discourse studies it appears that A arguments are more frequently topical, i.e. filled by referents higher on the hierarchy, while O arguments are more frequently borne by NPs with referents lower on the hierarchy, especially with rhematic and new referents (DuBois et al. 2003, Jäger 2007, among others). Because more frequent patterns generally tend to be less marked, these findings from discourse patterns allow formulation of the following hypothesis:

(49) Hypothesis:

Higher-ranking As and lower-ranking Os are more likely to be assigned a zero-marked case form than lower-ranking As and higher-ranking Os, respectively.

A popular variant of this hypothesis (originating in Silverstein 1976) equates ‘zero-marked’ with what is a distributionally unmarked form, i.e., a nominative or absolutive case. Two predictions follow. First, we can predict that across languages, pronouns (which necessarily rank high), but not nouns (which vary in their ranking), prefer accusative over other non-neutral alignment, so that they are in the unmarked nominative when in the A function. However, comparing the data on pronoun and noun alignment patterns in an expanded version of Nichols’ (1992) genealogically-balanced sample, this prediction has only marginal statistical support (Fisher Exact Test, $p = .075$, $N = 197$).¹⁹ This contrasts with areal factors which do have highly significant effects on the distribution of alignment types (cf. Nichols 1992).

The second prediction is that, if there is a difference in alignment within the same language, higher-ranking arguments are expected to show nominative {S,A}-alignment (or no case) and lower-ranking arguments to show absolutive {S,O}-alignment (or no case), while the reverse is unexpected. Comrie (2005) tests this prediction for the difference between pronouns and lexical nouns and finds a 20:3 support. However, the number of relevant languages (i.e. with a difference in markedness between pronouns and nouns) is small ($N = 23$), and genealogical and areal patterns are again a possible confounding variable.²⁰ Another prediction of (49) is that higher-ranking referents in O function tend to align with G arguments with overt

¹⁹ Cf. www.uni-leipzig.de/~autotyp. I removed cases of splits within categories to get clearer signals. The prediction has no good statistical support in Comrie’s (2005) dataset either.

²⁰ For example, of the 20 languages that support the hypothesis, Eskimo and Pama-Nyungan representatives might be oversampled in the database; Comrie also notes that 4 of the supporting languages (i.e. 20%) are from Australia.

(dative) case marking while the more common lower-ranking O arguments tend to align with less-marked T or S arguments. This has not been systematically tested.

In summary, despite its popularity, the statistical evidence for referential hierarchy effects on case alignment is weak. Historical aspects relating to descent and contact appear to be just as relevant. Indeed, specific etymologies and paradigm structures are often demonstrably relevant. For example, if an ergative develops from an instrumental, a limitation of the ergative to low-ranking As is to be expected just because animate nouns may never have had instrumental form to begin with (Garrett 1990). Or, demonstratives often inflect following the same paradigm as lexical nouns, and they share the function of introducing new referents in discourse, unlike personal pronouns. Under such conditions, we would expect the distribution of case alignment to follow part-of-speech categories (and their functions in discourse), and less directly the semantic notions of the referential hierarchies. Similarly, a third person pronoun may have lexical noun etymology and thereby inherit its case-paradigm, leading to a split between SAP and third person governed by paradigm structure rather than semantics. Or, an ergative system might survive in pronouns while lexical NPs lose case or develop new accusative marking, and this might result in a distribution that reverses what is predicted by discourse frequency (such as happened in a number of Dardic and other Indo-Aryan languages: Filimonova 2005).

6.2 Constructional effects on GR distributions

Going back to early proposals by Anderson (1976), another popular idea is that some constructions universally favor {S,A} relations while others are more flexible. In general, behavioral constructions (as defined in Section 5) are claimed to favor {S,A}, while coding constructions (especially case constructions) are expected to balance the odds for {S,A} vs. {S,O} more evenly. This idea also underlies early notions of ‘deep’ vs. ‘surface’ (or ‘syntactic’ vs. ‘morphological’) ergativity (e.g. Comrie 1978; Dixon 1994): many languages have ‘surface’ ergativity only, i.e. {S,O} alignments in their coding constructions but not in their behavioral constructions, or at least not all of them. Languages with ‘deep’ ergativity, i.e. with {S,O}-relations in behavioral constructions, appear to be less common.

Some theories propose in addition that among the behavioral constructions, those involving control, imperatives or reflexives universally favor accusative alignment. This is sometimes even claimed to be an absolute condition (e.g., Dixon 1994; Manning 1996, among others), but at least for control constructions, there are counterexamples (see the data in Section 4.5), and imperatives and reflexives often do not reference a syntactic GR notion to begin with (Section 4.9). Still, it is possible that these three construction types indeed have a probabilistic preference for accusative alignments. At present, there are no sufficiently rich databases to empirically evaluate this possibility.

One constructional effect that can be investigated is whether case and agreement construction differ in their preference for various GR types. Working on different samples, Siewierska (2004) and Haspelmath (2005) both observe a significant preference for accusative over other non-neutral alignment in agreement as opposed to

case constructions (where there is no clear preference),²¹ and for {O,T}-alignments in case as opposed to agreement constructions.

While these are preferences across languages, patterns of preference have also been proposed for the distribution of GRs across constructions within the same language. Elaborating on proposals by Kazenin (1994), Croft (2003) and others, it is likely that there is a hierarchy of GR constructions along the following lines:

- (50) case > agreement > relativization / focus / operator floating > conjunction
reduction > coreference constructions / coreference marking

The hypothesis then is that ergatively-aligned GRs in lower-ranking constructions in a language increase the odds for such GRs in higher-ranking constructions in the same language. However, as there are as yet no sufficiently rich typological databases on GRs in behavioral constructions, this hypothesis cannot be empirically tested and must remain speculative for a while. What is clear is that there are no absolute laws here: there are languages with accusatively-aligned relative constructions (active participles) but ergatively-aligned coreference constructions (e.g. Belhare, discussed above); and there are languages with accusative-aligned case but ergatively-aligned relative (e.g. Oirata, discussed above), quantifier floating (Japanese; Donohue 2007) or agreement (Siewierska 2004: 54) constructions. And it was noted long ago that even so thoroughly {S,A}-oriented languages as the Indo-European languages of Europe, align the S argument of at least a lexical subset of intransitive verbs (called ‘unaccusative’ since Perlmutter 1978) with the O argument, leading to traces of ergatively-aligned relative and other constructions (e.g. with past participle relativization). This all confirms the point made at the outset of this chapter: GRs hold of constructions, and not of languages. (In other words, once-popular expressions like ‘ergative language’ are simply senseless.)

7. Conclusions and prospects for further research

Research over the past three decades has been largely driven by the distributional theories discussed above, and this has spawned much descriptive research and has led to the discovery of the great diversity of GRs as we now know it. However, there has been a strong focus in this research on the specific distribution of {S,A} vs. {S,O} relations, and this has been at the expense of other relations. For example, research on GRs in behavioral constructions tends to neglect reference-based relations like the proximative and obviative GR, but it could very well be that there are interesting and significant trends in the distribution of such GRs as well. Likewise, GRs appear to distribute very unevenly over split intransitives, but there have not been many typological studies of this (though now see Donohue & Wichmann 2007).

More generally, given the large variable space, it is virtually impossible to estimate *a priori* which values on which variables will reveal significant clusters worldwide. Focusing on just one or two values of one single variable (i.e. on {S,A} vs {S,O} in the role subset variable) might mask other interesting distributional patterns. And finally, as noted earlier, the actual distribution of GR patterns reflects areal

²¹ This can be confirmed by an expanded version of Nichols’ (1992) sample ($N = 233$): the odds for {S, A}-relations in agreement rules are 2.55 times higher than in case rules (Fisher Exact Test, $p = .001$). In a 2 (case vs. agreement) \times 4 (macrocontinents) logistic regression model, the areal factor also reaches significance at a .05 level, but there is no significant interaction.

factors, and a proper understanding of frequency distributions needs to factor in not only linguistic variables like the ones in Table 3 but also historical information about language and population movements (cf. Nichols 1992). In short, GR typology has much work ahead here, and many interesting patterns are yet to be discovered once a multivariate approach is taken.

Meanwhile, the main reason why we lack large databases on GRs in behavioral constructions is that detailed descriptions of GRs have become standard in reference grammars only over the past two decades. And much more is still needed. A general message that can be drawn from a typological point of view is that the most informative descriptions do not ponder at length whether or not the language has or has not a subject (which is a theoretically dubious question anyway: Dryer 1997). What is more informative is to describe each GR-sensitive construction in the language and to describe in detail how the GRs in it are defined, and to what kinds of information they are sensitive. The variables described in this chapter are meant to help in this by providing a toolkit for comparing GRs across constructions in a single language, just as well as across languages.

Further reading

The great typological diversity in GRs both within and across languages has been the topic of intensive research since the 1970s, with a particularly strong impact from the collections of papers in Li (1976) and Plank (1979) and the work on pivots by Foley & Van Valin (1984). The nature and extent of {S,O} relations is the topic of a classic monograph by Dixon (1994). Hierarchical GR choice in the Americas is surveyed by Zúñiga (2006); for an in-depth study on Tagalog, see Kroeger (1993). For GR systems with lexical class conditions, see Donohue and Wichmann's (2007) state-of-the-art collection of papers. The construction-specific nature of GRs has been a great challenge for formal models of grammar. See Farrell (2005) for a recent survey of various responses.

The quantitative distribution of GRs across constructions and referents has recently received increased attention, e.g., Comrie (2005), Siewierska (2003; 2004), Haspelmath (2005), or Bickel 2007. Nichols (1992) offers extensive statistical analyses of the distribution of GRs in case and agreement constructions across word order and morphological complexity, in addition to a thorough discussion of the areal and historical principles that explain much of the current distributions. (On word order and alignment specifically, also see Siewierska 1996).

Acknowledgments

I am indebted to Martin Haspelmath, Jae Jung Song, Alena Witzlack, and an anonymous reviewer for helpful comments on earlier draft. I am alone responsible for all remaining unclarities and mistakes.

References

- Anderson, S. R. (1976), 'On the notion of subject in ergative languages', in C. N. Li (ed.), *Subject and topic*, New York: Academic Press, 1 – 23.
- Angdembe, T. M. (1998), 'Antipassive via noun incorporation: future of the Limbu object agreement', *Journal of Nepalese Studies*, 2, 17 – 25.
- Baker, M. C. (1996), *The Polysynthesis Parameter*, New York: Oxford University Press.
- Bickel, B. (1995), 'In the vestibule of meaning: transitivity inversion as a morphological phenomenon', *Studies in Language*, 19, 73 – 127.
- (1999), 'From ergativus absolutus to topic marking in Kiranti: a typological perspective', *Proceedings of the 25th Annual Meeting of the Berkeley Linguistics Society*, 38 – 49.
- (2004a), 'Hidden syntax in Belhare', in A. Saxena (ed.), *Himalayan languages: past and present*, Berlin: Mouton de Gruyter, 141 – 90.
- (2004b), 'The syntax of experiencers in the Himalayas', in P. Bhaskararao and K. V. Subbarao (eds.), *Non-nominative subjects*, Amsterdam: Benjamins, 77 – 112.
- (2006), 'Clause-level vs. predicate-level linking', in I. Bornkessel, et al. (eds.), *Semantic role universals and argument linking: theoretical, typological and psycholinguistic perspectives*, Berlin: Mouton de Gruyter, 155 – 90.
- (2007), 'On the scope of the referential hierarchy in the typology of grammatical relations.' in G. G. Corbett and M. Noonan (eds.), *Case and grammatical relations. Papers in honor of Bernard Comrie*, Oxford: Oxford University Press.
- Bickel, B. and Yāḍava, Y. P. (2000), 'A fresh look at grammatical relations in Indo-Aryan', *Lingua*, 110, 343 – 73.
- Bickel, B. and Nichols, J. (2001), 'Syntactic ergativity in light verb complements', *Proceedings of the 27th Annual Meeting of the Berkeley Linguistics Society*, ## – ##.
- Bickel, B. and Gaenszle, M. (2005), '1P antipassives and the political history of the Southern Kirant', Paper presented at the 11th Himalayan Languages Symposium, Bangkok, December 6 - 9, 2005; handout available at http://www.uni-leipzig.de/~ff/cpdp/frameset_publ.html.
- Bickel, B. and Nichols, J. (2007), 'Inflectional morphology', in T. Shopen (ed.), *Language typology and syntactic description*, Cambridge: Cambridge University Press (Revised second edition).
- (in press), 'Case-marking and alignment', in A. Malchukov and A. Spencer (eds.), *The Handbook of Case*, Oxford: Oxford University Press, ## – ##.
- Bickel, B., Gaenszle, M., Rai, A., Rai, P. D., Rai, S. K., Rai, V. S., and Sharma(Gautam), N. P. (2007), 'Two ways of suspending object agreement in Puma: between incorporation, antipassivization, and optional agreement', *Himalayan linguistics*, 7, 1 - 18.
- Bossong, G. (1985), *Empirische Universalienforschung: differentielle Objektmarkierung in neuiranischen Sprachen*, Tübingen: Narr.
- Bresnan, J., Cueni, A., Nikitina, T., and Baayen, R. H. (2004), 'Predicting the dative alternation', Paper presented at the KNAW Academy Colloquium on the Cognitive Foundations of interpretation; Ms., Stanford University.
- Clark, T. W. (1989), *Introduction to Nepali*, Kathmandu: Ratna Pustak Bhandar.
- Comrie, B. (1975), 'The anti-ergative: Finland's answer to Basque', *Proceedings of the 11th Annual Meeting of the Chicago Linguistics Society*, 112 – 21.
- (1978), 'Ergativity', in W. P. Lehmann (ed.), *Syntactic typology: studies in the phenomenology of language*, Austin: University of Texas Press.
- (1981), *Language universals and linguistic typology*, Chicago: University of Chicago Press.
- (1988), 'Coreference and conjunction reduction in grammar and discourse', in J. A. Hawkins (ed.), *Explaining language universals*, London: Blackwell, 186 – 208.
- (2005), 'Alignment of case marking', in M. Haspelmath, et al. (eds.), *The world atlas of language structures*, Oxford: Oxford University Press, 398 – 405.
- Croft, W. (2003), *Typology and universals*, Cambridge: Cambridge University Press [2nd edition].
- DeLancey, S. (1981), 'An interpretation of split ergativity and related patterns', *Language*, 57, 626 – 57.
- Dixon, R. M. W. (1972), *The Dyirbal language of North Queensland*, Cambridge: Cambridge University Press.
- (1979a), 'Ergativity', *Language*, 55, 59 – 138.

- (1979b), 'Corrections and comments concerning Heath's "Is Dyirbal ergative?"', *Linguistics*, 17, 1003 – 15.
- (1994), *Ergativity*, Cambridge: Cambridge University Press.
- (1995), 'Complement clauses and complementation strategies', in F. R. Palmer (ed.), *Grammar and meaning: essays in honor of Sir John Lyons*, Cambridge: Cambridge University Press, 175 – 220.
- Donohue, M. (2007), 'Semantic alignment systems: what's what and what's not', in M. Donohue and S. Wichmann (eds.), *Semantic alignment: typological and descriptive studies*, Oxford: Oxford University Press, ## – ##.
- Donohue, M. and Brown, L. (1999), 'Ergativity: some additions from Indonesia', *Australian Journal of Linguistics*, 19, 57 – 76.
- Donohue, M. and Wichmann, S. (eds.) (2007), *Semantic alignment: typological and descriptive studies*, Oxford: Oxford University Press.
- Dowty, D. R. (1991), 'Thematic proto-roles and argument selection', *Language*, 67, 547 – 619.
- Dryer, M. S. (1997), 'Are grammatical relations universal?' in J. Bybee, J. Haiman, and S. A. Thompson (eds.), *Essays on language function and language type dedicated to T. Givón*, Amsterdam: Benjamins, 117 – 43.
- Du Feu, V. (1996), *Rapanui*, London: Routledge.
- DuBois, J. W. (1987), 'The discourse basis of ergativity', *Language*, 63, 805 – 55.
- DuBois, J. W., Kumpf, L. E., and Ashby, W. J. (eds.) (2003), *Preferred argument structure: grammar as architecture for function*, Amsterdam: Benjamins.
- Dunn, M. J. (1999), *A grammar of Chukchi*, Ph. D. dissertation, Australian National University, Canberra.
- Durie, M. (1985), *A grammar of Acehnese*, Dordrecht: Foris.
- (1987), 'Grammatical relations in Acehnese', *Studies in Language*, 11, 365 – 99.
- Evans, N. (1995), *A grammar of Kayardild*, Berlin: Mouton de Gruyter.
- (1997), 'Role or cast? Noun incorporation and complex predicates in Mayali', in A. Alsina, J. Bresnan, and P. Sells (eds.), *Complex predicates*, Stanford: CSLI, 397 – 430.
- Faarlund, J. T. (1998), 'Symétrie et dissymétrie des actants centraux', in J. Feuillet (ed.), *Actance et valence dans les langues de l'Europe*, Berlin: Mouton de Gruyter, 147 – 92.
- Farrell, P. (2005), *Grammatical relations*, Oxford: Oxford University Press.
- Farrell, P., Marlett, S. A., and Perlmutter, D. M. (1991), 'Notions of subjecthood and switch reference: evidence from Seri', *Linguistic Inquiry*, 22, 431 – 56.
- Foley, W. A. (1986), *The Papuan languages of New Guinea*, Cambridge: Cambridge University Press.
- (1998), 'Symmetrical voice system and precategoriality in Philippine languages', Paper presented at the Workshop on Voice and Grammatical Functions in Austronesian Languages, The University of Queensland, July 1, 1998. Ms. available at <http://www.sultry.arts.usyd.edu.au/LFG98/austro/workshop.htm>.
- Foley, W. A. and Van Valin, R. D. (1984), *Functional syntax and universal grammar*, Cambridge: Cambridge University Press.
- Fox, B. (1987), 'The noun phrase accessibility hierarchy reinterpreted: subject primacy or the absolutive hypothesis?' *Language*, 63, 856 – 70.
- Garrett, A. (1990), 'The origin of NP split ergativity', *Language*, 66, 261 – 96.
- Gerdts, D. B. (1988), *Object and absolutive in Halkomelem Salish*, New York: Garland.
- Givón, T. (2001), *Syntax*, Amsterdam: Benjamins.
- Hagège, C. (1974), 'Les pronoms logophoriques', *Bulletin de la Société de Linguistique de Paris*, 69, 287 – 310.
- Hale, K. (1983), 'Warlpiri and the grammar of nonconfigurational languages', *Natural Language and Linguistic Theory*, 1, 5 – 47.
- Haspelmath, M. (2005), 'Argument marking in ditransitive alignment types', *Linguistic Discovery* [<http://linguistic-discovery.dartmouth.edu>], 3.
- Haude, K. (2006), *A grammar of Movima*, Ph. D. dissertation, Radboud University, Nijmegen.
- Henderson, J. (1995), *Phonology and grammar of Yele, Papua New Guinea*, Canberra: Pacific Linguistics.
- Huisman, R. D. (1973), 'Angaatha verb morphology', *Linguistics*, 110, 43 – 54.
- Jäger, G. (2007), 'Evolutionary Game Theory and typology: a case study', *Language*, 83, 74 – 109.
- Kazenin, K. I. (1994), 'Split syntactic ergativity: toward an implicational hierarchy', *Sprachtypologie und Universalienforschung*, 47, 78 – 98.
- Keenan, E. (1976), 'Towards a universal definition of "subject"', in C. N. Li (ed.), *Subject and topic*, New York: Academic Press, 305 – 33.
- Kiparsky, P. (2002), 'On the architecture of Pāṇini's grammar', Ms., Stanford University, www.stanford.edu/~kiparsky.
- Kroeger, P. (1993), *Phrase structure and grammatical relations in Tagalog*, Stanford: CSLI.

- LaPolla, R. J. (1993), 'Arguments against 'subject' and 'object' as viable concepts in Chinese', *Bulletin of the Institute of History and Philology, Academia Sinica*, 63, 759 – 813.
- Levinson, S. C. (2003), 'Syntactic ergativity in Yéll Dnye, the Papuan language of Rossel Island', Ms. Max-Planck-Institute for Psycholinguistics, Nijmegen.
- Li, C. N. (ed.), (1976), *Subject and Topic*, New York: Academic.
- MacDonald, L. (1990), *A grammar of Tauya*, Berlin: Mouton de Gruyter.
- Manning, C. D. (1996), *Ergativity: argument structure and grammatical relations*, Stanford: CSLI.
- Merlan, F. (1985), 'Split intransitivity: functional oppositions in inflections', in J. Nichols and A. C. Woodbury (eds.), *Grammar inside and outside the clause: some approaches to theory from the field*, Cambridge: Cambridge University Press, 324 – 62.
- Mithun, M. (1999), *The languages of Native North America*, Cambridge: Cambridge University Press.
- Moravcsik, E. (1978), 'On the distribution of ergative and accusative patterns', *Lingua*, 45, 233 – 79.
- Munro, P. and Gordon, L. (1982), 'Syntactic relations in Western Muskogean', *Language*, 58, 81 – 115.
- Nagano, Y. (1984), *A historical study of the rGyarong verb system*, Tokyo: Seishido.
- Nichols, J. (1992), *Language diversity in space and time*, Chicago: The University of Chicago Press.
- (2007), 'Why are stative-active languages rare in Eurasia? Typological perspectives on split subject marking', in M. Donohue and S. Wichmann (eds.), *Semantic alignment: typological and descriptive studies*, Oxford: Oxford University Press, ## – ##.
- Perlmutter, D. M. (1978), 'Impersonal passives and the unaccusative hypothesis', *Proceedings of the 4th Annual Meeting of the Berkeley Linguistics Society*, 157 – 89.
- Peterson, J. (1998), *Grammatical relations in Pali and the emergence of ergativity in Indo-Aryan*, München: LINCOM EUROPA.
- Pilhofer, G. (1933), *Grammatik der Kâte-Sprache in Neuguinea*, Berlin: Reimer.
- Plank, F. (ed.), (1979), *Ergativity*, New York: Academic Press.
- Pokharel, M. (2054), *Nepālī Vākya Vyākaraṇ, Kāṭhmāṇḍū: Nepāl Rājakīya Pragya-Pratiṣṭhān*.
- Polinsky, M. and Potsdam, E. (2006), 'Expanding the scope of control and raising', *Syntax*, 9, ## – ##.
- Primus, B. (1999), *Cases and thematic roles*, Tübingen: Niemeyer.
- Reed, I., et al. (1977), *Yup'ik Eskimo grammar*, Fairbanks: Alaska Native Language Center, University of Alaska.
- Reesink, G. P. (1983), 'Switch reference and topicality hierarchies', *Studies in Language*, 7, 215 – 46.
- Reis, M. (1982), 'Zum Subjektbegriff im Deutschen', in W. Abraham (ed.), *Satzglieder im Deutschen: Vorschläge zur syntaktischen, semantischen und pragmatischen Fundierung*: Tübingen, 171 – 211.
- Rhodes, R. (1976), *The morphosyntax of the Central Ojibwa verb*, University of Michigan, Ann Arbor.
- (1994), 'Valency, inversion, and thematic alignment in Ojibwe', *Proceedings of the 20th Annual Meeting of the Berkeley Linguistics Society*, 431 – 46.
- Rigsby, B. and Rude, N. (1996), 'Sketch of Sahaptin, a Sahaptian language', in I. Goddard (ed.), *Handbook of North American Indians, Vol. 17: Languages*, Washington: Smithsonian Institution, 666 – 92.
- Roberts, J. R. (1988), 'Amele switch-reference and the theory of grammar', *Linguistic Inquiry*, 19, 45 – 64.
- Robins, R. H. (1958), *The Yurok language: grammar, texts, lexicon*, Berkeley: University of California Press.
- Schultze-Berndt, E. and Himmelmann, N. (2004), 'Depictive secondary predicates in cross-linguistic perspective', *Linguistic Typology*, 8, 59 – 131.
- Scott, G. (1978), *The Fore language of Papua New Guinea*, Canberra: Pacific Linguistics (PL-B47).
- Seidl, A. and Dimitriadis, A. (1997), 'The discourse function of object marking in Swahili', *Proceedings of the 33rd Annual Meeting of the Chicago Linguistics Society*, 373 – 89.
- Siewierska, A. (1996), 'Word order type and alignment', *Sprachtypologie und Universalienforschung*, 49, 149 – 76.
- (2003), 'Person agreement and the determination of alignment', *Transactions of the Philological Society*, 101, 339 – 70.
- (2004), *Person*, Cambridge: Cambridge University Press.
- Silverstein, M. (1976), 'Hierarchy of features and ergativity', in R. M. W. Dixon (ed.), *Grammatical Categories in Australian Languages*, New Jersey: Humanities Press, 112 – 71.
- Simpson, J. (1991), *Warlpiri morpho-syntax: a lexicalist approach*, Dordrecht: Kluwer.
- Speas, M. (1990), *Phrase structure in natural language*, Dordrecht: Kluwer.
- Timberlake, A. (1980), 'Oblique control of Russian reflexivization', in C. V. Chvany and R. D. Brecht (eds.), *Morphosyntax in Slavic*, Columbus, Ohio: Slavica, 235 – 59.
- Van Valin, R. D. (1981), 'Grammatical relations in ergative languages', *Studies in Language*, 5, 361 – 94.
- (2005), *Exploring the syntax-semantics interface*, Cambridge: Cambridge University Press.
- Van Valin, R. D. and LaPolla, R. J. (1997), *Syntax: structure, meaning, and function*, Cambridge: Cambridge University Press.
- Verhoeven, E. (2005), *Experiential constructions in Yucatec Maya*, Ph.D. dissertation, Universität Erfurt.

- Watkins, L. J. and McKenzie, P. (1984), *A grammar of Kiowa*, Lincoln: University of Nebraska Press.
- Zúñiga, F. (2006), *Deixis and alignment: inverse systems in indigenous languages of the Americas*, Amsterdam: Benjamins.